

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



David Hauzar

Zpracování digitálních snímků videokymografických záznamů jako podpůrný nástroj pro diagnostiku hlasivek

Ústav teorie informace a automatizace Akademie věd České
republiky

Vedoucí diplomové práce: Prof. Ing. Jan Flusser, DrSc.

Studijní program: Informatika, softwarové systémy

2010

Děkuji Prof. Ing. Janu Flussrovi, DrSc., mému školiteli, za jeho podněty a cenné připomínky k mé diplomové práci. Vřelé díky patří RNDr. Barbaře Zitové, PhD. za obětavou podporu při psaní diplomové práce, za její cenné rady, návrhy a velké množství poznámek. V neposlední řadě děkuji RNDr. Janu Švecovi, Ph.D. et Ph.D a Mgr. A. Vetešníkov, Ph.D. za jejich práci na kterou jsem mohl navázat. Můj velký dík patří všem mým učitelům MFF UK za jejich vedení a péči během celého studia.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

Praha 9. srpna 2010

David Hauzar

Obsah

1	Úvod	7
1.1	Hlasivky a vznik hlasu	7
1.2	Vyšetření hlasivek	7
2	Cíle práce	10
3	Analýza problému	11
3.1	Potlačení šumu ve snímku	11
3.2	Segmentace snímku	16
3.2.1	Detekce hran	17
3.2.2	Deformovatelné modely	17
3.2.3	Prahování	20
3.2.4	Segmentace založená na oblastech	23
3.3	Vibrační rysy hlasivek	24
3.4	Cíle detekce	33
3.5	Vyhodnocování výsledků	34
4	Popis řešení	35
4.1	Možnosti extrakce základních příznaků	35
4.2	Základní algoritmus a jeho konfigurace	36
4.3	Hledání prahu pro segmentaci hlasivkové štěrby	38
4.4	Extrakce základních příznaků ze snímku	40
4.5	Odvozené charakteristiky kmitání	47
5	Vyhodnocení	49
6	Závěr a otevřené problémy	59
A	Obsah CD	60
B	Uživatelská příručka	61
B.1	Instalace	61
B.2	Spuštění programu	61
B.3	Vizualizace příznaků ve snímku	62
B.4	Zadání výchozích a známých hodnot základních příznaků	63
B.5	Vyhodnocení výsledků	64

C Architektura programu	68
Literatura	72

Název práce: Zpracování digitálních snímků videokymografických záznamů jako podpůrný nástroj pro diagnostiku hlasivek

Autor: David Hauzar

Katedra (ústav): Ústav teorie informace a automatizace Akademie věd České republiky

Vedoucí diplomové práce: Prof. Ing. Jan Flusser, DrSc.

e-mail vedoucího: flusser@utia.cas.cz

Abstrakt: Videokymografie je metoda umožňující zachytit řadu charakteristik kmitání hlasivek důležitých pro stanovení jejich diagnózy. Manuální extrakce těchto charakteristik je v řadě případů problematická, automatická extrakce by stanovení diagnózy usnadnila a zpřesnila. Automatické zpracování snímků komplikuje především velké množství šumu na snímcích a dále velká variabilita průběhu kmitání hlasivek různých pacientů. Cílem práce je vybrat charakteristiky kmitání hlasivek vhodné pro automatickou extrakci a jejich extrakci implementovat.

S ohledem na specifické vlastnosti videokymografických snímků byla vyvinuta implementace automatické extrakce, která může sloužit jako základ nástroje pro podporu kymografické diagnostiky hlasivek. Vytvořený systém pro vyhodnocení výsledků automatické extrakce byl spolu se sadou snímků a sadou charakteristik kmitání extrahovaných manuálně použit k podrobnému testování implementace. Klíčová slova: videokymografie, vyšetření hlasivek, digitální zpracování obrazu, segmentace obrazu

Title: Analysis of digital videokymographic images as a tool for vocal folds diagnostics

Author: David Hauzar

Department: Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic

Advisor: Prof. Ing. Jan Flusser, DrSc.

Advisor's e-mail address: flusser@utia.cas.cz

Abstract: Videokymography is a medical imaging method of revealing and diagnosing vocal cords vibrations in voice disorders. Manual data extraction is problematic while automatic extraction can facilitate and refine the diagnostic process. However, automatic processing is hampered by signal noise and considerable variability in vocal cords vibrations of individual patients. The objective of the present study is to identify typical characteristics of vocal cords vibrations suitable for automatic extraction and diagnostic interpretation and to implement such automatic extraction.

The automatic extraction tool that was implemented reflects specific features of videokymographic images. The system for interpretation of automatic extraction results was developed and tested against manually extracted vibrations data and images. The tool can support kymographic diagnosis of vocal cords

disorders.

Keywords: videokymography, vocal chord diagnosis, digital image processing, image segmentation

Kapitola 1

Úvod

Videokymografie je metoda umožňující zaznamenat pohyb hlasivek a zobrazit některé jeho charakteristiky v digitálním snímku. Manuální určení některých charakteristik pohybu hlasivek je obtížné a nepřesné. Digitální zpracování obrazu poskytuje techniky, které umožňují řadu těchto charakteristik učit automaticky. Tak je možné vytvořit nástroj, který poslouží k podpoře kymografické diagnostiky hlasivek.

1.1 Hlasivky a vznik hlasu

Hlasivky jsou základem hlasového ústrojí. Funkce hlasového ústrojí je vytvářet základní hlas, jehož dalšími úpravami vzniká lidská řeč. Dle [1] jsou hlasivky tvořeny dvěma hlasovými valy pokrytými sliznicí umístěnými v horní části hrtanu. Mezi nimi se vytváří hlasivková štěrbina (glottis) trojúhelníkového tvaru.

Vznik hlasu není zcela objasněn. Dle [1] vzniká hlas nejpravděpodobněji dle myoelastické teorie. Tedy činností hlasivkových svalů a vlivem tlaku vydechovaného vzduchu. Při sevření hlasivek pod nimi vzniká přetlak a nad nimi naopak podtlak. Přetlak pod hlasivkami vede k jejich oddálení. Kvantum vzduchu pod hlasivkami unikne a vzniknou podmínky pro jejich nové přiblížení. Jestliže je proud vzduchu plynulý, hlasivky kmitají.

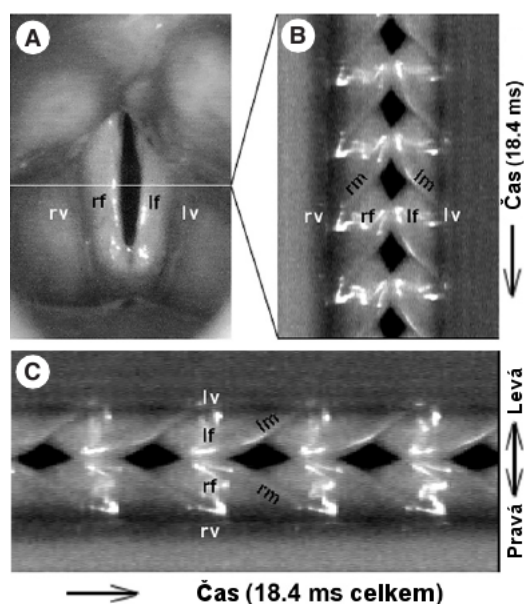
1.2 Vyšetření hlasivek

Frekvence kmitání hlasivek se běžně pohybuje v rozmezí 150 až 400 Hz. Při zpěvu může dosáhnout hodnot blížících se 900 Hz. To znamená, že kamery snímající dle současně používaných televizních standardů 25/30 snímků za sekundu, nemožnou zachytit kmity hlasivek.

K překonání tohoto problému je v klinické praxi používána *stroboskopie*. Vibrující hlasivky jsou osvětlovány blikajícím zdrojem světla. Při synchronizaci světelných záblesků s kmity hlasivek se hlasivky ve snímaném záznamu jeví jako by byly v klidu. Při vhodné desynchronizaci jsou hlasivky osvětleny vždy v jiné fázi kmitu a tím dojde ke zdánlivému zpomalení kmitů. To umožňuje pozorovat

kmity hlasivek i za pomoci “pomalých” kamer pracujících s televizním standardem. Použití stroboskopie je však omezeno pouze na záznamy periodických kmitů. V případě nepravidelného kmitání hlasivek poskytuje stroboskopie zcela zkreslený obraz vibrací.

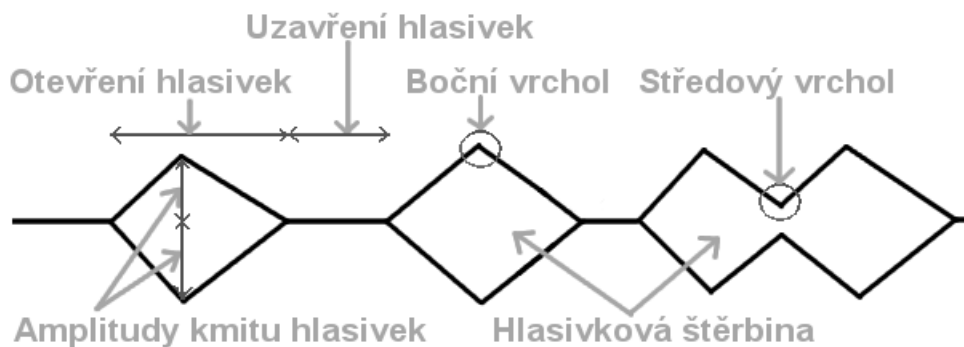
Jako alternativa byla RNDr. Janem Švecem, Ph.D. et Ph.D. navržena *videokymografická kamera* [3]. Ta pracuje ve dvou módech. Ve standardním módu kamera snímá, stejně jako standardní televizní kamera, snímky s rozlišením 768x576 (720x486) pixelů a s frekvencí 25/30 snímků za sekundu. Ve videokymografickém nebo vysokorychlostním módu kamera snímá pouze jeden řádek (vrchní) a tím dosahuje frekvence 7812,5 (7875) řádků za sekundu s rozlišením 768 (720) pixelů. Videokymografický snímek je pak dvourozměrný obraz složený z řádků sejmutých videokymografickou kamerou tak, že řádek sejmutý kamerou v daném časovém okamžiku je ve snímku pod řádkem sejmutým v předchozím časovém okamžiku. Příklad snímku z videokymografické kamery je na obrázku 1.1.



Obrázek 1.1: Snímek z videokymografické kamery. Zdravé hrdlo ženy v průběhu fonace na laryngoskopickém (A) a videokymografickém (B, C) snímku. Videokymografické snímky zobrazují vibrace hlasivek v pozici, která je označena čarou A. Videokymomogram B je zobrazen v původní orientaci s časovou osou směřující dolů. Videokymogram C je identický jako B, ale je rotovaný tak, že časová osa směřuje horizontálně doleva. rf/lf - pravá/levá hlasivka; rv/lm - pravá/levá ventrikulární řasa; rm/lm - pravá/levá slizniční vlna postupující bočně přes hlasivky. Obrázek a popis obrázku převzat z [6]

Na obrázku 1.2 je schematicky zakreslený videokymografický snímek spolu s popisem jeho částí důležitých při charakterizaci kmitů a určení diagnózy. Mezi důležité charakteristiky kmitání zachycené ve videokymografickém snímku patří například absence kmitání hlasivek, frekvence kmitání, poměr doby uzavření k době otevření hlasivek, charakteristiky průběhu kmitání nebo charakteristiky va-

riability průběhu po sobě jdoucích vibrací. Podrobnější rozbor těchto charakteristik následuje v dalších kapitolách.



Obrázek 1.2: Schéma videokymografického snímku.

Pro přesný záznam vibrací hlasivek bez omezení na periodicitu kmitů je možné v dnešní době použít také *vysokorychlostní kamery*. Díky současnému technickému pokroku se tyto kamery staly dostupnější. Jejich uplatnění v klinické praxi je však stále omezené. Mezi hlavní nevýhody vysokorychlostních kamer patří nadprodukce informace. Určení užitečných charakteristik kmitání z videosekvence pohybu hlasivek je problematické.

Kapitola 2

Cíle práce

Hlavním cílem práce je navrhnout a implementovat automatickou extrakci vybraných charakteristik kmitání hlasivek z videokymografických snímků. Z analýzy používaných metod budou zvoleny charakteristiky kmitání hlasivek vhodné pro automatickou extrakci a extrakce těchto charakteristik bude navržena a implementována.

Videokymografie umožňuje zachytit pohyb hlasivek ve dvourozměrném snímku. Zpracování videokymografických snímků komplikuje především velké množstvím šumu na snímcích a dále velká odlišnost jednotlivých videokymografických snímků způsobená odlišnými vlastnosmi hlasivek jednotlivých pacientů.

Návrh automatické extrakce zohlední specifické vlastnosti videokymografických snímků. Implementace bude testována na reálných datech. Vzhledem ke zmíněné vysoké odlišnosti snímků je důležité testovat implementaci na dostatečně reprezentativním souboru videokymografických snímků, proto bude vytvořen nástroj umožňující úspěšnost automatické extrakce vyhodnotit.

Kapitola 3

Analýza problému

Aby bylo možné ve snímku určit příznaky vhodné k automatické detekci, je nutné zkoumat charakteristiky pohybu hlasivek, které jsou patné z videokymografického snímku. Pro extrakci příznaků je vhodné oddělit hlasivkovou šterbinu od jejího okolí. K tomu se nabízí využití techniky segmentace. Aplikace některých technik segmentace bude v této práci zkoumána. Při snímání videokymografickou kamerou vzniká na snímcích velké množství šumu, který je nutno kvůli dalšímu zpracování potlačit.

3.1 Potlačení šumu ve snímku

Množství šumu ve snímku je ovlivněno například intenzitou osvětlení snímaného objektu nebo teplotou snímače. Příklad videokymografického snímku s přítomným šumem je na obrázku 3.1. Intenzitu pixelů v pořízeném snímku g můžeme vyjádřit vztahem:

$$g(x, y) = f(x, y) + \eta(x, y)$$

kde člen $f(x, y)$ obsahuje intenzity ve snímané scéně a člen $\eta(x, y)$ obsahuje hodnoty aditivního šumu.

Existuje velké množství metod umožňujících eliminovat vliv šumu. *Lokální metody* počítají intenzitu ve výsledném snímku na základě okolí příslušného pixelu ve vstupním snímku. Mezi tyto metody patří například lokální průměrování nebo mediánový filtr. *Globální metody* naopak při stanovení intenzity pixelu pracují s celým snímek. Mezi tyto metody patří například metody založené na waveletové transformaci. Dále je uveden stručný výčet některých základních metod:

- **Průměrování snímků**

Pokud je k dispozici více vzájemně na sebe registrovaných snímků stejné, statické scény, je možno vytvořit snímek s menším množstvím šumu tak, že



Obrázek 3.1: Výřez videokymografického snímku s patrným šumem.

intenzita každého pixelu ve výsledném snímku bude rovná průměru intenzit pixelů ve vstupních snímcích. Tedy:

$$g(i, j) = \frac{1}{m} \sum_{k=1}^m f_k(i, j)$$

kde f_1, \dots, f_m jsou vzájemně na sebe registrované snímky stejné, statické scény a g je výsledný snímek s potlačeným šumem. Tato metoda využívá předpokladu, že šum je nezávislá veličina s nulovou střední hodnotou. Lze dokázat, že náhodná veličina vzniklá jako průměr n stejně rozdělených nezávislých náhodných veličin o daném rozptylu σ^2 , má rozptyl $\frac{\sigma^2}{n}$. Na základě toho dojde ke snížení rozptylu šumu ve výsledném snímku. V případě videokymografie scéna není statická a proto není možno tuto metodu použít.

• Lokální průměrování snímku

Metoda lokálního průměrování snímku využívá stejně jako metoda průměrování snímků výše toho, že průměrování náhodné veličiny snižuje její rozptyl. Intenzita pixelu ve výsledném snímku se spočítá jako průměr intenzit pixelů v okolí pixelu v původním snímku. Tedy:

$$g(x, y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s, t)$$

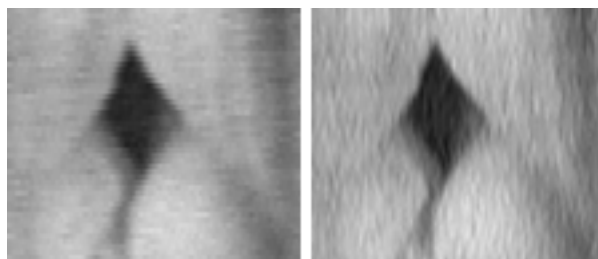
kde S_{xy} reprezentuje množinu souřadnic velikosti $m \times n$ v obdélníkovém okolí bodu (x, y) v tomto bodu centrovaném. Nevýhoda této metody je rozmazání hran, které nastává v důsledku průměrování. Pomocí volby tvaru okolí je ale možné povahu rozmazání ovlivnit. Pokud má například okolí více řádků než sloupců, dojde k většímu rozmazání ve vertikálním směru než ve směru horizontálním.

Příklad použití lokálního průměrování je na obrázku 3.2. Při použití menšího okolí dojde k potlačení šumu a lehkému rozmazání hran, při použití většího okolí dojde k většímu potlačení šumu, ale také k výraznějšímu rozmazání hran. Se zřetelem na další použití je na obrázku 3.3 demonstrován efekt použití nečtvercového okolí. V prvním případě má okolí pouze jednu

řádku a dochází tak k rozmazání v horizontálním směru, ve druhém případě má okolí pouze jeden sloupec a dochází k rozmazání ve vertikálním směru. V případě rozmazání v horizontálním směru dochází k zúžení hlasivkové štěrbině v místě otevření a uzavření, v případě rozmazání ve vertikálním směru k zúžení hlasivkové štěrbině v místech amplitud.



Obrázek 3.2: Lokální průměrování snímku z obrázku 3.1. Použití okolí 3×3 (vlevo) a okolí 5×5 (vpravo).



Obrázek 3.3: Lokální průměrování snímku z obrázku 3.1. Použití okolí 1×5 (vlevo) a okolí 5×1 (vpravo).

- **Mediánový filtr**

Intenzita pixelu ve výstupním snímku je při použití mediánového filtru stanovena jako medián hodnot z okolí S_{xy} odpovídajícího pixelu ve vstupním snímku. Tedy:

$$g(x, y) = \underset{(s, t) \in S_{xy}}{\text{median}}(g(s, t))$$

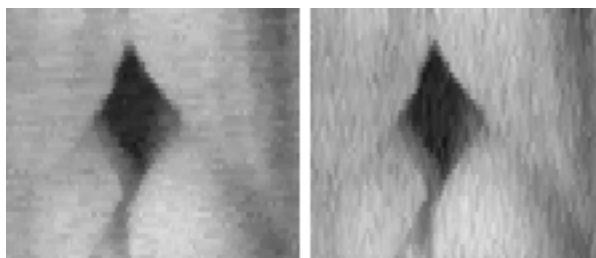
Výhoda mediánového filtru oproti průměrování je, že při stejné velikosti okolí dochází k menšímu rozmazání hran. Mediánový filtr vykazuje výborné výsledky především při potlačení impulzivního šumu. Nevýhoda tohoto filtru je ztráta pozvolných jasových přechodů a detailů v obraze a zaoblení ostrých hran objektů. Povahu rozmazání lze stejně tak jako v případě lokálního průměrování ovlivnit volbou tvaru okolí.

Příklad použití mediánového filtru na výřez snímku hlasivky je na obrázku 3.4. Je vidět, že dochází k znatelně menšímu rozmazání hran než v případě

lokálního průměrování. Při použití většího okolí dochází k viditelné ztrátě detailů - tento efekt je možné pozorovat u původně ostrých hran, které jsou v důsledku mediánového filtru zaobleny. Stejně jako v případě lokálního průměrování je s ohledem na další použití metody demonstrován efekt nečtvercového okolí. Snímky na obrázku 3.5 vznikly za použití nečtvercového okolí s pouze jednou řádkou respektive jedním sloupcem. Podobně jako v případě lokálního průměrování je efekt použití nečtvercového okolí nejvíce patrný v oblastech otevření a uzavření hlasivek a v oblastech amplitud kmitů.



Obrázek 3.4: Mediánový filtr použitý na snímek z obrázku 3.1. Použití okolí 3×3 (vlevo) a okolí 5×5 (vpravo).



Obrázek 3.5: Mediánový filtr použitý na snímek z obrázku 3.1. Použití okolí 1×5 (vlevo) a okolí 5×1 (vpravo).

• Adaptivní filtry

Dosud diskutované metody potlačení šumu pracují na celém snímku stejně, bez ohledu na rozdílné charakteristiky aktuálně zpracovávané části snímku. Chování adaptivních filtrů se naproti tomu liší na základě charakteristiky okolí aktuálně zpracovávaného pixelu. Lze tak dosáhnout nižšího rozmazání v oblastech hran a většího vyhlazení v homogenních oblastech.

Mezi známé adaptivní filtry patří *metoda rotující masky*. Tato metoda potlačuje šum pomocí průměrování pixelů v daném okolí. Aby se zamezilo rozmazání hran, k průměrování dochází pouze v nejvíce homogenní oblasti aktuálně zpracovávaného okolí. Filtr pracuje tak, že okolí právě zpracovávaného pixelu se rozdělí na několik podokolí. V každém se spočítá rozptyl. Dále se spočítá průměr okolí s nejnižší hodnotou rozptylu a ten se dosadí

jako hodnota pixelu ve výsledném snímku. Výhoda tohoto filtru je, že nedochází k rozmazání snímku, filtr má dokonce mírně ostřicí efekt. Nevýhoda je, že dochází k ztrátě detailů.

Mezi další adaptivní filtry patří *adaptivní mediánový filtr*. V průběhu filtrování dochází ke změně velikosti okolí S_{xy} právě zpracovávaného pixelu. Za použití následujícího značení:

$$\begin{aligned} z_{min} &= \text{minimální hodnota intenzity v } S_{xy} \\ z_{max} &= \text{maximální hodnota intenzity v } S_{xy} \\ z_{med} &= \text{medián hodnot intenzit v } S_{xy} \\ z_{xy} &= \text{hodnota intenzity pixelu na souřadnici } (x, y) \\ S_{max} &= \text{maximální velikost } S_{xy} \end{aligned}$$

je podle [7] možno adaptivní mediánový popsat takto:

Fáze A: $A1 = z_{med} - z_{min}$
 $A2 = z_{med} - z_{max}$
 Pokud $A1 > 0$ a $A2 < 0$ jdi do fáze B
 Jinak zvětši velikost okolí
 Pokud velikost okolí $\leq S_{max}$ opakuj fázi A
 Jinak vrať z_{med}

Fáze B: $B1 = z_{xy} - z_{min}$
 $B2 = z_{xy} - z_{max}$
 Pokud $B1 > 0$ a $B2 < 0$ vrať z_{xy}
 Jinak vrať z_{med}

Algoritmus předpokládá, že hodnoty z_{min} a z_{max} představují hodnoty impulsního šumu. První podmínka fáze A tedy testuje, zda je hodnota z_{med} impulsní šum. Pokud není, jde se do fáze B. Pokud je, dojde k zvětšení okolí. Ve fázi B se pak analogicky otestuje, zda je aktuálně zpracovávaný bod impulsní šum. Pokud není, vrátí algoritmus původní hodnotu pixelu, pokud je, vrátí algoritmus stejně jako klasický mediánový filtr medián z aktuálního okolí.

Adaptivní mediánový filtr v porovnání s mediánovým filtrem lépe zachovává detaily. Je to jednak díky tomu, že pracuje s proměnnou velikostí okna a medián se nepočítá ze zbytečně velkého okolí a také díky tomu, že mediánem se zpracovávaný pixel nahrazuje jenom v případě, že algoritmus detekuje, že jde o šum. Adaptivní mediánový filtr také umožňuje potlačit impulsní šum s větším výskytem než klasický mediánový filtr, potlačuje i neimpulsivní šum a umožňuje lépe než metoda rotující masky použít obdélníkové okolí s pouze jedním řádkem nebo sloupcem a zachovat tak struktury ve vertikálním nebo horizontálním směru.

Příklad použití adaptivního mediánového filtru na potlačení šumu ve výřezu videokymografického snímku je na obrázku 3.6. Po jedné iteraci algoritmu

téměř nedošlo k rozmazání hran. Na druhou stranu potlačení šumu není tak výrazné jako u ostatních zmíněných metod. Po více iteracích algoritmu je výsledek podobný jako u jedné iterace klasického mediánového filtru. S ohledem na další použití metody je na obrázku 3.7 demonstrován efekt použití nečtvercového okolí s pouze jedním sloupcem. Po více iteracích dochází k velmi dobrému potlačení šumu při velmi dobrém zachování hran tvořících začátek a konec otevření hlasivkové štěrbiny. Naopak hrany amplitud jsou silně rozmazané. Na obrázku 3.8 je vidět efekt použití nečtvercového okolí s pouze jednou řádkou. Stejně jako u filtru s jedním sloupcem dochází po více iteracích je šum velmi dobře potlačen. Dochází ale naopak k velmi dobrém zachování hran amplitud a silnému rozmazání hran tvořících začátek a konec otevření hlasivkové štěrbiny.



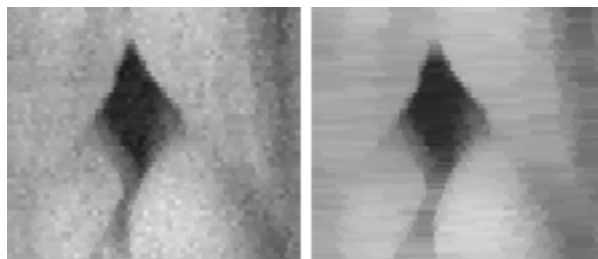
Obrázek 3.6: Adaptivní mediánový filtr použitý na snímek z obrázku 3.1. Použití okolí velikosti 3×3 až 15×15 (vlevo) a použití čtyř iterací tohoto algoritmu se stejnou velikostí okolí (vpravo).



Obrázek 3.7: Adaptivní mediánový filtr použitý na snímek z obrázku 3.1. Použití okolí velikosti 3×1 až 15×1 (vlevo) a použití čtyř iterací tohoto algoritmu se stejnou velikostí okolí (vpravo).

3.2 Segmentace snímku

Další důležitou úlohou, kterou je potřeba řešit během zpracování videokymografického snímku, je segmentace, neboli rozdělení obrazu na konzistentní oblasti nebo objekty. Videokymografický snímek se dá rozdělit do dvou oblastí. První oblast tvoří hlasivková štěrbina a druhou oblast tvoří její okolí. Na obrázku 3.9-a je videokymografický snímek, kde jsou tyto oblasti označeny. Následuje přehled některých základních segmentačních metod a ukázky jejich aplikace na videokymografické snímky.



Obrázek 3.8: Adaptivní mediánový filtr použitý na snímek z obrázku 3.1. Použití okolí velikosti 3×1 až 15×1 (vlevo) a použití čtyř iterací tohoto algoritmu se stejnou velikostí okolí (vpravo).

3.2.1 Detekce hran

Tyto metody detekují hrany, které od sebe oddělují segmenty na obrázku. Hrany jsou místa, kde v obrázku dochází k rychlému změnu jasu. K nalezení těchto míst je možné použít první nebo druhé derivace obrazové funkce. V praxi se využívá operátorů, které první nebo druhou derivaci aproximují.

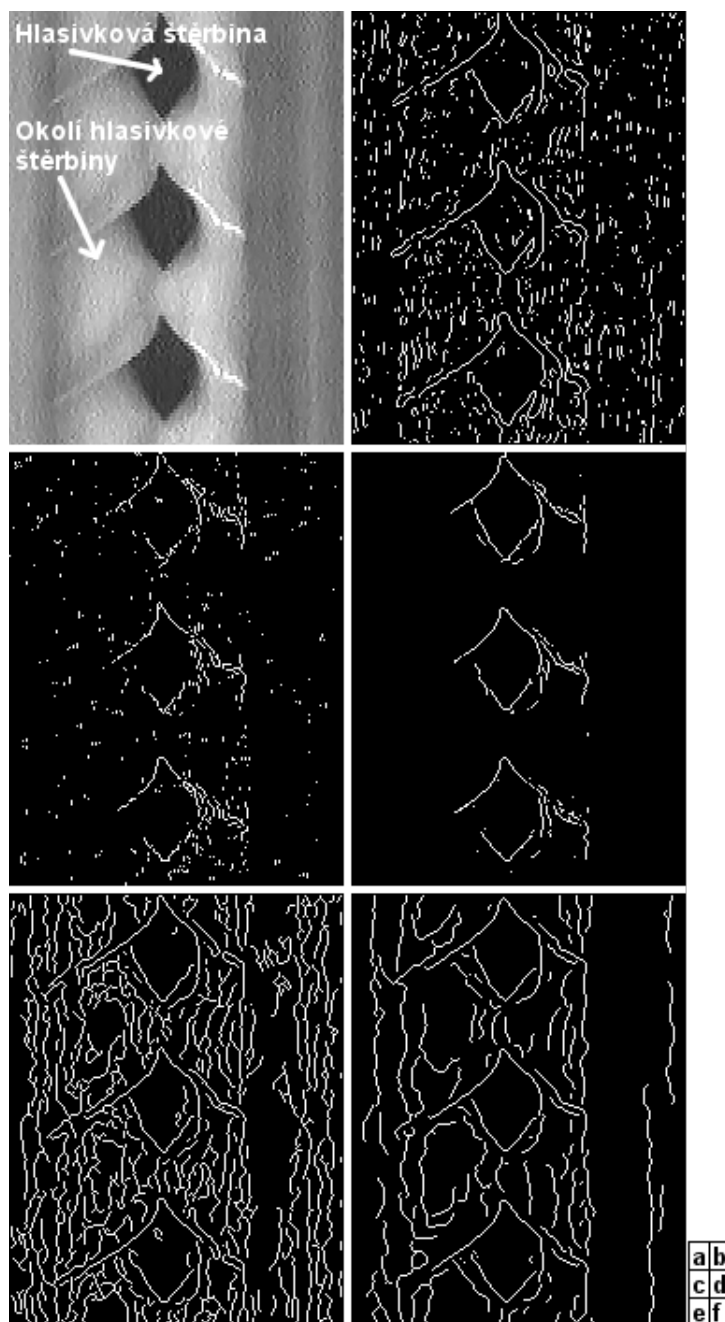
K aproximaci první derivace se používá například *Sobelův operátor*, *Robinsonův operátor* nebo *Prewittův operátor*. Často používaný je *Cannyho detektor*. Ten vyhladí obrázek pomocí Gaussova filtru, spočítá aproximaci gradientu například pomocí Sobelova operátoru a potlačí hodnoty gradientu, které nejsou lokální maxima. Nakonec provede prahování takto upraveného gradientu daným prahem a napojování hodnot gradientu, které jsou menší než tento práh, ale větší než jiný daný práh. Druhou derivaci je možno aproximovat například *Laplaceho operátorem*. Výhoda tohoto operátoru je invariance vůči otočení, nevýhoda pak větší citlivost vůči šumu plynoucí z vlastností druhé derivace. Pro snížení citlivosti vůči šumu je často nutné vyhladit obrazovou funkci. Za tímto účelem lze použít Gausián, aplikací Laplaceova operátoru na Gausiánu vzniká takzvaný *LoG* operátor.

Výsledky jednotlivých detektorů hran jsou na obrázku 3.9. Laplaceův operátor je v porovnání s ostatními operátory citlivější na šum. Je to kvůli použití aproximace druhé derivace. Díky invarianci vůči otočení lépe než Sobelův operátor detekuje hrany různých směrů. Sobelův operátor je méně citlivý na šum, ale nalezené hrany jsou značně nespojité. Při použití vyhlazení šumu lokálním průměrováním se spojitost nalezených hran zlepší. Hrany nalezené Cannyho detektorem jsou v porovnání s ostatními detektory nejvíce spojitě.

Problém všech detektorů je to, že nalezené hrany jsou nespojitě, tedy zcela neohraničují hlasivkovou šterbinu a z poskytnutých výsledků tak segment odpovídající hlasivkové šterbině nelze přímo určit.

3.2.2 Deformovatelné modely

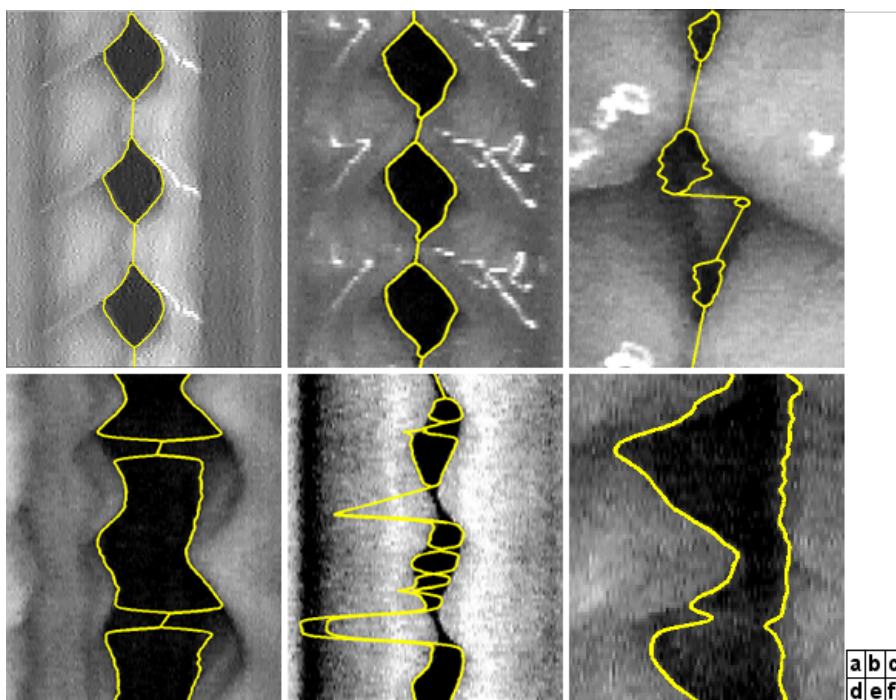
Deformovatelné modely, jinak také hadi (*snakes*) [2], jsou křivky nebo povrchy určující hranice mezi jednotlivými segmenty tím, že se deformují na základě působení vnitřních a vnějších sil tak, aby minimalizovaly svou celkovou energii. Vni-



Obrázek 3.9: Výřez videokymografického snímku hlasivek (a), výsledek aplikace Laplaceova operátoru (b), výsledek aplikace Sobelova operátoru (c), výsledek aplikace Sobelova operátoru na snímek s potlačeným šumem pomocí mediánového filtru s velikostí okolí 5×5 (d), výsledek použití Cannyho filtru (e), výsledek použití Cannyho filtru na snímek s potlačeným šumem pomocí mediánového filtru s velikostí okolí 5×5 (f).

třní síly charakterizují napětí uvnitř křivky a udržují křivku hladkou. Vnější síly udržují křivku u hran obrázku, často jsou definované pomocí gradientu obrazové funkce.

Mgr. A. Vetešníkem, Ph.D. byla pomocí této metody implementována segmentace hlasivkové štěrbiny [8]. Na obrázku 3.10 jsou příklady použití této metody na videokymografické snímky. Na rozdíl od detekce hran nalezne tato metoda spojitě ohraničení segmentů. Je ale problematické nastavit parametry metody tak, aby co křivka co nejlépe kopírovala hranice segmentů a aby neodbíhala po falešných hranicích. Tato metoda také kvůli požadavku na určitou hladkost křivky vyhlazuje hrany. Toto vyhlazování je patrné zejména v oblastech amplitud kmitů, může ale dojít i k vyhlazení vlnění průběhu vibrace, dvojitého středového vrcholu a středové nehladkosti. Při snížení požadavku na hladkost křivky a menším vyhlazení ostrých hran se tato metoda špatně vyrovnává s velkým množstvím šumu přítomným ve snímcích a interferencí s okolím. Vznikají artefakty na konturách snímků, které mohou být s výše zmíněnými poruchami průběhu vibrací naopak zaměněny. Především vlivem šumu někdy dojde k zcela chybnému směřování křivky.



Obrázek 3.10: Segmentace videokymografických snímků pomocí deformovatelných modelů provedená pomocí nástroje vytvořeného Mgr. A. Vetešníkem, Ph.D. [8]. Dobře segmentovaný snímek (a), nadměrné vyhlazení hran snímku (b), špatná segmentace vlivem interference hlasivky s okolím (c), špatná segmentace vlivem málo ostrých hranic (d), špatná segmentace vlivem šumu (e), artefakty na konturách snímků zaměnitelné s poruchami průběhu vibrací (f).

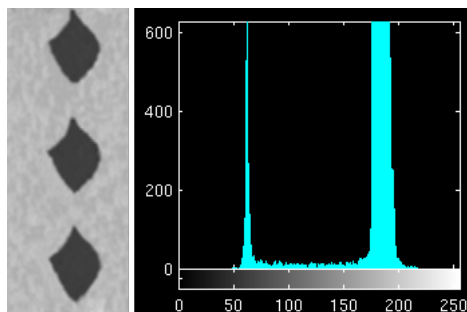
3.2.3 Prahování

Prahování je metoda segmentace založená na předpokladu, že objekty v jednotlivých oblastech jsou charakterizovány danou odrazivostí povrchů a pixely příslušející do dané oblasti tak mají intenzitu v daném rozsahu intenzit. Důležitý je požadavek, aby se intenzity pixelů v jednotlivých oblastech nepřekrývaly. Základní prahování detekuje v obraze dva segmenty, pozadí a popředí. Formálně je to transformace vstupního obrazu f na výstupní segmentovaný binární obraz g podle vztahu:

$$y = \begin{cases} 1 & \text{pro } f(x, y) \geq T_{xy} \\ 0 & \text{pro } f(x, y) < T_{xy} \end{cases} \quad (3.1)$$

Pixely příslušející segmentům odpovídajícím objektům ve výstupním obrazu jsou označeny hodnotou 1, pixely příslušející pozadí jsou označeny hodnotou 0. T_{xy} . Pokud je hodnota *prahu* T_{xy} konstantní, mluvíme o *globálním prahování*, pokud se mění v závislosti na souřadnicích (x, y) , jedná se o *lokální prahování*.

Protože oblast hlasivkové štěrbině je tmavá a oblast odpovídající okolí světlá, za ideálních podmínek nenastávají při použití prahování žádné problémy. Na obrázku 3.11 je vidět ideální snímek hlasivek a jeho histogram. Na histogramu snímku jsou vidět dva výrazné vrcholy dobře oddělené výrazným údolím. Vrchol histogramu vlevo odpovídá pixelům hlasivkové štěrbině a vrchol histogramu vpravo odpovídá pixelům okolí, jako práh je možné zvolit hodnotu odpovídající některému sloupci histogramu ležící v údolí mezi těmito vrcholy.

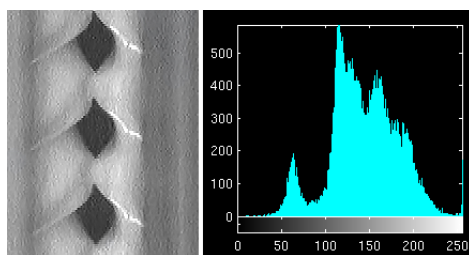


Obrázek 3.11: “Ideální“ snímek hlasivek a jeho histogram.

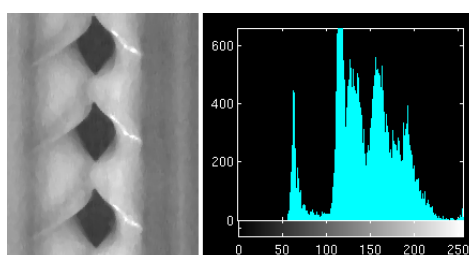
V reálném videokymografickém snímku jsou vrcholy histogramu odpovídající hlasivkové štěrbině méně výrazné. Je to jednak díky tomu, že zejména intenzita okolí není zcela uniformní a okolí tak obsahuje i tmavší pixely a také díky přítomnosti šumu ve snímku. Šum jednak ztěžuje nalezení hodnoty prahu a také může způsobit přiřazení pixelů zasažených šumem do nesprávné oblasti. Potlačení šumu je proto velmi důležité.

Příklad reálného videokymografického snímku a jeho histogramu je na obrázku 3.12. Na obrázku 3.13 je tentýž snímek s potlačeným šumem a jeho histogram. Vrcholy histogramu jsou v tomto případě výraznější. Na obrázku 3.14 je videokymografický snímek s tmavými pruhy v okolí hlasivkové štěrbině, jeho histogram a prahovaný snímek. Tmavé pruhy mají nižší intenzitu než některé pixely

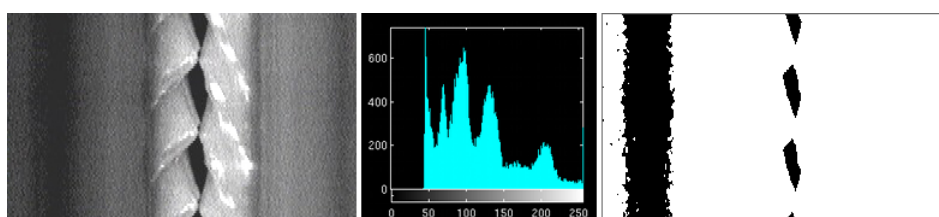
hlasivkové štěrby a jsou proto v prahovaném snímku chybně klasifikovány do segmentu patřící hlasivkové štěrbině.



Obrázek 3.12: Reálný snímek hlasivek a jeho histogram.



Obrázek 3.13: Snímek hlasivek s potlačeným šumem a jeho histogram.

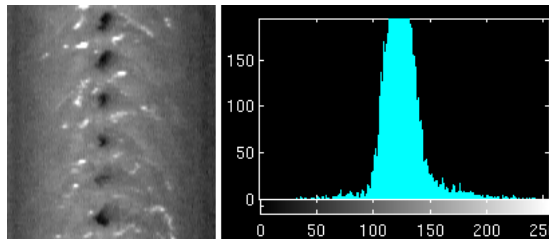


Obrázek 3.14: Videokymografický snímek s tmavými pruhy v okolí hlasivkové štěrby, jeho histogram a oprahovaný snímek.

Další problém nastává, pokud je oblast příslušející jednomu segmentu výrazně menší než oblast příslušející druhému segmentu. Protože do ní patří málo pixelů, je vrchol vytvořený z pixelů tohoto segmentu malý a práh je tak obtížně určitelný. Příklad takového videokymografického snímku je na obrázku 3.15. Řešení je pracovat s výřezem snímku, kde hlasivková štěrbina zabírá větší relativní část nebo vytvořit histogram pouze z částí ležících v blízkosti hran. Tvar histogramu se pak stane méně závislý na relativní velikosti segmentů.

Základní algoritmus hledání prahu

V [7] byl navržen jednoduchý algoritmus hledání prahu. Vychází z předpokladu, že snímek obsahuje dvě třídy objektů - objekty popředí a pozadí a tyto dvě třídy tvoří v histogramu výrazné vrcholy oddělené dostatečně hlubokým údolím. Algoritmus je možné popsat takto:



Obrázek 3.15: Videokymografický snímek kde je oblast tvořící hlasivkovou šterbinu relativně malá vzhledem k oblasti tvořící okolí.

1. Zvol iniciální odhad globálního prahu T . Tento odhad musí být větší než minimální hodnota intenzity v obrázku a menší než maximální hodnota intenzity v obrázku. Dobrý iniciální odhad je například průměrná hodnota intenzity v obrázku.
2. Segmentuj obrázek za použití prahu T . Vzniknou dvě skupiny pixelů. Skupina pixelů G_1 bude obsahovat pixely s hodnotami intenzit $> T$ a G_2 pixely s hodnotami intenzit $\leq T$.
3. Spočítej průměrnou hodnotu intenzit m_1 ve skupině G_1 a m_2 ve skupině G_2 .
4. Spočítej novou hodnotu prahu:

$$T = \frac{1}{2}(m_1 + m_2) \quad (3.2)$$

5. Opakuj kroky 2 až 4 dokud je rozdíl mezi hodnotou T v po sobě následujících iteracích algoritmu větší než ΔT .

Hledání optimálního prahu Otsuovou metodou

Další možnost, jak stanovit hodnotu prahu je Otsuova metoda popsaná například v [7]. Stejně jako algoritmus výše tato metoda předpokládá, že snímek obsahuje dvě třídy objektů které v histogramu tvoří dva vrcholy. Otsuova metoda hledá práh o hodnotě k tak, aby rozdělení do výsledných segmentů maximalizovalo variabilitu mezi třídami σ_B^2 :

$$\sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2 \quad (3.3)$$

kde:

$$P_1 = \sum_{q=0}^{k-1} p_q(r_q)$$

$$P_2 = \sum_{q=k}^{L-1} p_q(r_q)$$

$$\begin{aligned}
m_1 &= \sum_{q=0}^{k-1} q \frac{p_q(r_q)}{P_1} \\
m_2 &= \sum_{q=k}^{L-1} q \frac{p_q(r_q)}{P_2} \\
m_G &= \sum_{q=0}^{L-1} q p_q(r_q) \\
p_r(r_q) &= \frac{n_q}{n}
\end{aligned}$$

kde L je počet intenzit v obrázku, $q = 0, 1, 2, \dots, L - 1$, n je počet pixelů v obrázku a n_q je počet pixelů s intenzitou r_q . Otsuovu metodu lze snadno implementovat tak, že se hodnota σ_B^2 spočte pro všechny možné hodnoty prahu k_q a vybere se práh, který dává největší hodnotu σ_B^2 .

Na obrázku 3.16 nahoře je ilustrováno použití hledání prahu pomocí Otsuovy metody na videokymografický snímek. Je vidět, že použití této metody na celý snímek není příliš vhodné. Intenzita pixelů v okolí hlasivkové štěrby není homogenní a pixely okolí proto v histogramu tvoří několik vrcholů. Je tak porušen předpoklad, že histogram obsahuje dva dobře oddělené vrcholy. Okolí hlasivkové štěrby navíc obsahuje podstatně větší množství pixelů než hlasivková štěrbina a Otsuova metoda nenajde práh v údolí histogramu mezi vrcholem tvořeným pixely hlasivkové štěrby a následujícím vrcholem. Na obrázku 3.16 nahoře dole je ilustrováno použití hledání prahu pomocí Otsuovy metody ve výřezu snímku okolo středu kmitání hlasivkové štěrby. V rámci tohoto výřezu je okolí hlasivkové štěrby více homogenní a také počet pixelů náležejících do hlasivkové štěrby je vzhledem k počtu pixelů náležejících do jejího okolí větší. Nalezený práh je proto lepší.

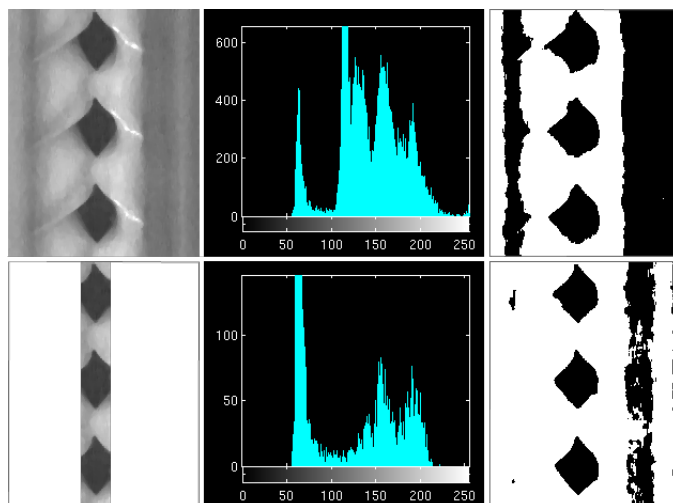
3.2.4 Segmentace založená na oblastech

Segmentační metody patřící do této skupiny konstruují oblasti přímo. Snaží se rozčlenit obraz do maximálních souvislých oblastí tak, aby byly z daného hlediska homogenní. Mezi tyto metody patří například *narůstání oblastí*, algoritmus *štěpení a slučování oblastí* nebo algoritmus *rozvodí* (*watershed*). Dále je pro představen jeden zástupce tohoto přístupu.

Narůstání oblastí

Narůstání oblastí seskupuje pixely nebo podoblasti do větších oblastí na základě definovaných kritérií narůstání. K iniciální množině pixelů, takzvaným "semínkům", se přidávají pixely s nimi sousedící splňující daná kritéria. Algoritmus končí, když neexistují žádné pixely sousedící s nějakou oblastí, které splňují tato kritéria.

Výběr semínek závisí na povaze řešeného problému. V případě videokymografických snímků je často na základě detekce předchozích parametrů známá oblast,



Obrázek 3.16: Videokymografický snímek s potlačeným šumem, jeho histogram a obraz získaný prahováním prahem vypočteným za použití Otsuovy metody (nahore), výřez videokymografického snímku, jeho histogram a obraz získaný prahováním snímku nahore prahem vypočteným za použití Otsuovy metody z výřezu snímku (dole).

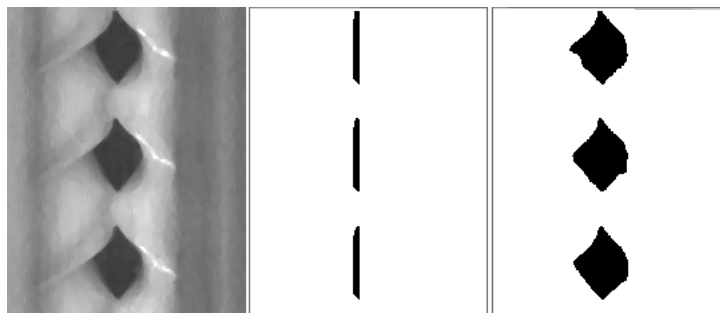
kde se určitě nachází hlasivková štěrbin. Například pokud je známá souřadnice středu vibrací, je možné semínka získat prahováním úzkého výřezu snímku kolem středu vibrací.

Kritérium podobnosti může zahrnovat deskriptory založené na intenzitě pixelů. Příklad takového kritéria je požadavek, aby velikost rozdílu intenzity pixelu ležícího v oblasti a sousedního pixelu připojovaného k této oblasti byla menší než daná hodnota. Další příklad jednoduchého kritéria je požadavek, aby intenzita připojovaného pixelu byla menší než daný práh. Existuje velké množství jiných kritérií, některé z nich jsou popsány v knize [7].

Na obrázku 3.17 je vidět použití metody narůstání oblastí pro segmentaci videokymografického snímku. Semínka tvoří pixely s intenzitou nižší než daný práh vzdálené od středu snímku nejvýše 3 pixely a kritérium podobnosti je požadavek, že intenzita pixelu musí být nižší než stejný práh. Práh je, stejně jako práh na obrázku 3.16 dole, určen pomocí Otsuovy metody aplikované na výřez snímku okolo středu vibrací. Narůstání oblastí na rozdíl od prahování správně segmentuje tmavé pruhy u krajů snímku do segmentu okolí hlasivky.

3.3 Vibrační rysy hlasivek

Videokymografickou kamerou bylo v pražském Centru péče o nemocné s poruchami hlasu, řeči a vadami sluchu MEDICAL HEALTHCOM v období 1996-2006 pořízeno okolo 7000 videokymografických nahrávek pokrývajících široké spektrum onemocnění hlasivek. Tato světově jedinečná sbírka byla použita pro vymezení charakteristických rysů vibrací hlasivek zobrazených ve videokymografických zá-



Obrázek 3.17: Segmentace narůstáním oblastí. Videokymografický snímek (vlevo), semínka označená černými pixely (uprostřed), výsledek segmentace (vpravo).

znamech [6]. Byly navrženy základní kategorie vibračních rysů hlasivek, které mohou být použity při stanovení diagnózy. Tyto kategorie byly použity pro vytvoření dotazníku [4] sloužícího k vyhodnocení videokymografického snímku. Vyplňování dotazníku probíhalo zcela manuálně, bez automatické podpory. Tato kapitola podává přehled kategorií vibračních rysů hlasivek dle [6].

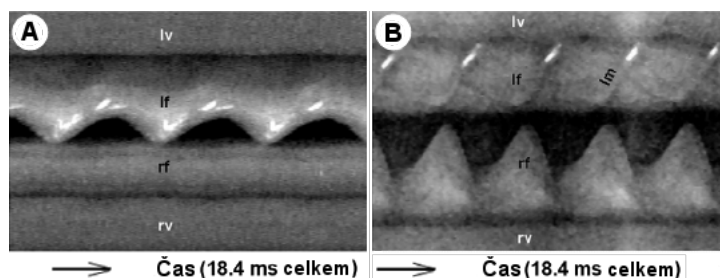
Obrázek 3.18: Dotazník pro vyhodnocení videokymografických snímků.

1. Absence kmitání hlasivek

Absence kmitání může být buď částečná nebo úplná a může nastat buď pouze u levé hlasivky, pouze u pravé hlasivky nebo u obou částí. Případy absence kmitání jsou znázorněny na obrázku 3.19. Dále se sleduje také počet kmitů obou částí.



Obrázek 3.19: Levá hlasivka kmitá, levá hlasivka kmitá částečně, levá hlasivka nekmitá. Převzato z [4]



Obrázek 3.20: **A)** Absence kmitání pravé hlasivky (rf) postižené nádorem. **B)** Částečná absence kmitání levé hlasivky (lf) zapříčiněná špatně umístěným silikonovým implantátem. Obrázek a popis obrázku převzat z [6].

Absence kmitání hlasivek je velmi vážný nález, který indikuje mimořádně tuhou tkáň hlasivky zabraňující hlasovému projevu. Může nastat například v důsledku zhoubného či nezhooubného nádoru nebo jizvy. Příklad snímku hlasivek, kdy došlo k absenci kmitání, je na obrázku 3.20.

2. Interference okolí s hlasivkami

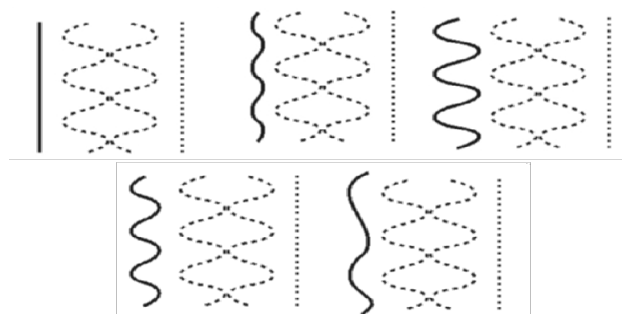
Může dojít jednak k vibraci hlasivek spolu s ventrikulárními řasami a také spolu s tekutinami. Sleduje se velikost amplitudy kmitání ventrikulárních řas v porovnání s amplitudou kmitání hlasivek 3.21-nahoře a frekvence kmitání ventrikulárních řas v porovnání s frekvencí kmitání hlasivek 3.21-dole. Dále se sleduje, zda dochází k interferenci kmitání tekutin a zda je vlivem interference snímek zkreslený a vyhodnocení snímku je tak problematické.

Interference okolí s hlasivkami může poškodit průběh vibrace hlasivek a tím ovlivnit výslednou kvalitu hlasu. Dále může výrazně ovlivnit zpracování snímku a je proto třeba s ní počítat. Příklad snímku, kdy došlo k interference okolí s hlasivkami, je na obrázku 3.22.

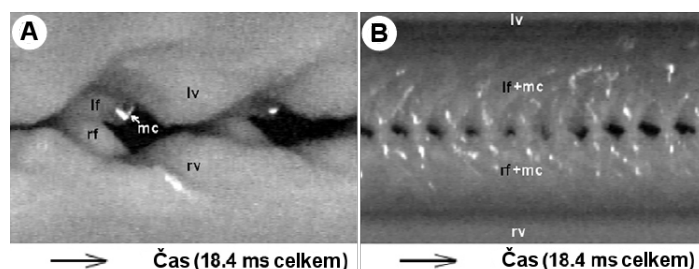
3. Doba uzavření hlasivkové štěrby

Sleduje se poměr délky otevření k délce uzavření hlasivkové štěrby. Důležité případy jsou znázorněny na obrázku 3.23.

Doba uzavření hlasivkové štěrby kratší než 20% a delší než 60% doby otevření hlasivek v místě největší amplitudy je považována za patologický nález. Na snímcích 3.20-B a 3.27-B nedochází k uzavření hlasivkové štěrby, na snímcích



Obrázek 3.21: Interference ventrikulárních řas. Velikost amplitudy kmitání ventrikulárních řas v porovnání s amplitudou kmitání hlasivek (nahore) a velikost frekvence kmitání ventrikulárních řas v porovnání s frekvencí kmitání hlasivek (dole). Převzato z [4]



Obrázek 3.22: **A)** Vibrace hlasivek spolu s ventrikulárními řasami (rv, lv) s tradiční diagnózou "hyperkinetická dysfonie". **B)** Hlen (mc) vibrující spolu s hlasivkami u ženy postižené zánětem hrtanu. Obrázek a popis obrázku převzat z [6].



Obrázek 3.23: Kategorie hodnoty poměru délky otevření k délce uzavření: nedochází k uzavření hlasivky, délka uzavření hlasivky je do 20% délky otevření hlasivky, 20-40%, 40-60%, více než 60%. Převzato z [4]

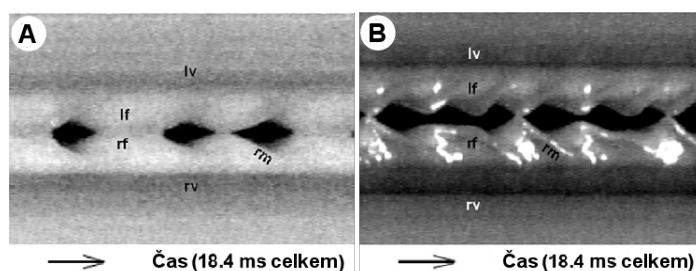
3.20-A a 3.25-B dochází k příliš krátkému zavření hlasivkové štěrbině a na snímku 3.29-A dochází k příliš dlouhému uzavření hlasivkové štěrbině.

4. Variabilita průběhu po sobě jdoucích vibrací

Po sobě jdoucí vibrace mohou vykazovat variabilitu délky trvání, amplitudy a variabilitu celkového tvaru vibrací. Mohou nastat dva případy variability vibrací. Synchronní variabilita nastává pokud se levá hlasivka chová stejně jako pravá hlasivka a asynchronní variabilita nastává v opačném případě. Důležité příklady variability průběhu po sobě jdoucích vibrací jsou znázorněny na obrázku 3.24.



Obrázek 3.24: Variabilita průběhu po sobě jdoucích vibrací levé hlasivky - nepatrná variabilita, střední a velká variabilita (vlevo), synchronní a asynchronní variabilita (vpravo). Převzato z [4]



Obrázek 3.25: **A)** Velká levo-pravá synchronní variabilita po sobě jdoucích vibrací hlasivek může s tradiční diagnózou "hyperkinetická dysfonie" **B)** Asynchronní variabilita po sobě jdoucích vibrací hlasivek ženy s cystou na levé hlasivce (cy). Obrázek a popis obrázku převzat z [6].

Variabilita průběhu po sobě jdoucích vibrací má za následek hrubý hlas. Asynchronní variabilita je znak odlišnosti levé a pravé hlasivky, synchronní variabilita je znak předozadní asymetrie. Příklad snímku hlasivek, kde dochází k variabilitě průběhu po sobě jdoucích vibrací je zobrazen na obrázku 3.25.

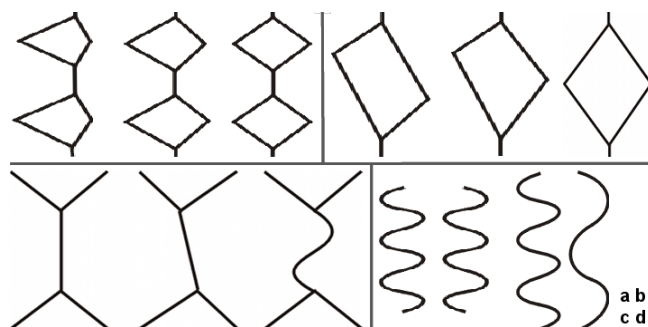
5. Levo-pravá asymetrie

Může nastat rozdíl v amplitudě kmitání levé a pravé části hlasivek, fázový rozdíl jednoho kmitu - tedy rychlejší otevírání jedné strany hlasivky než druhé a naopak rychlejší zavírání druhé než první. Dále posun osy hlasivkové štěrby při zavření a rozdíl frekvence kmitání levé a pravé hlasivky. Všechny tyto případy jsou znázorněny na obrázku 3.26.

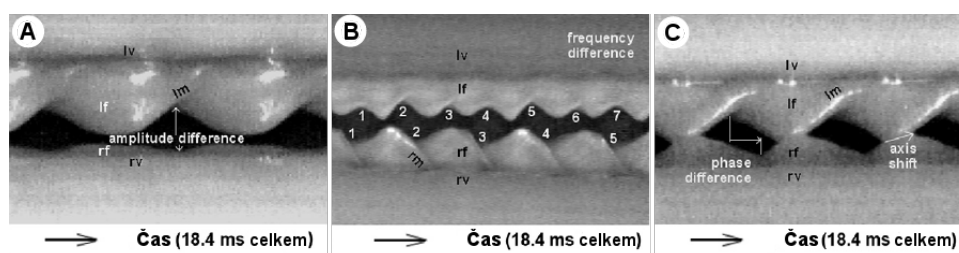
Očekává se, že fázový rozdíl je doprovázen posunutím osy. Malé asymetrie jsou normální, velké rozdíly mohou způsobit velké hlasové problémy. Příklad snímku s levo-pravo asymetrií je na obrázku 3.27.

6. Tvar bočních vrcholů

Boční vrcholy se nacházejí v místě středu amplitudy kmitů. Mohou být ostré nebo zakulacené 3.28. Větší zakulacení bočního vrcholu je projevem větší tuhosti sliznice, což je nepříznivý nález. Videokymografický snímek hlasivky s ostrým a neostrým bočním vrcholem je na obrázku 3.29.



Obrázek 3.26: Levo-pravá asymetrie. Rozdíl v amplitudě kmitání levé a pravé části hlasivek (a), fázový rozdíl jednoho kmitu (b), posun osy hlasivkové štěrbiny (c) a rozdíl frekvence levé a pravé hlasivky. Převzato z [4].



Obrázek 3.27: Levo-pravá asymetrie. **A)** Rozdíl amplitudy v případě pravostranného laryngeálního ochrnutí. **B)** Fázový rozdíl a posunutí osy v dalším případě pravostranného laryngeálního ochrnutí po tyreoplastice typu 1. **C)** Rozdíl ve frekvenci, tedy desynchronizace levé a pravé hlasivky v případě levostranného laryngeálního ochrnutí. Zatímco levá hlasivka vykonala sedm kmitů, pravá pouze pět kmitů. Obrázek a popis obrázku převzat z [6].

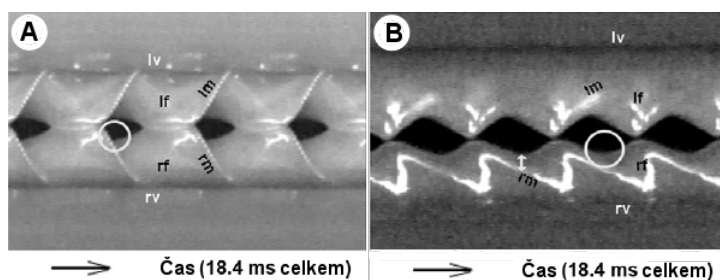


Obrázek 3.28: Kategorie tvaru bočních vrcholů levé hlasivky: ostré, spíše ostré, spíše kulaté, kulaté boční vrcholy. Převzato z [4].

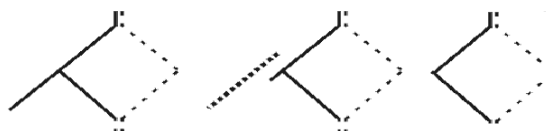
7. Bočně postupující slizniční vlny

Slizniční vlny mohou být na snímku zachyceny bez světelných odlesků nebo s odlesky 3.30, důležitý je dosah slizniční vlny 3.31 a v případě přítomnosti světelných odrazů také vzdálenost od amplitudy otevření hlasivkové štěrbiny 3.32.

Slizniční vlny patří mez sledované charakteristiky, významné jsou tyto případy: chybějící slizniční vlny, malé slizniční vlny, vzdálené slizniční vlny. Slizniční vlna je malá, pokud je menší než polovina velikosti hlasivek. Čím jsou slizniční vlny výraznější, tím ohebnější je sliznice. Snímek hlasivek s přítomnými slizničními vlnami je na obrázku 3.29.



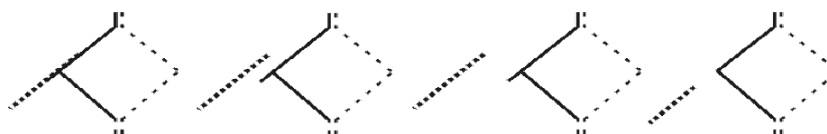
Obrázek 3.29: Tvar bočních vrcholů a rozdíly mezi slizničními vlnami. **A)** Ostrý boční vrchol (zakroužkován) ženy s normálně vypadajícími hlasivkami. Slizniční vlny (lm, rm) se táhnou bočně přes celé hlasivky a jsou zvýrazněny odlesky světla (bílé kontury) na vrcholech vln. **B)** Zakulacený boční vrchol (zakroužkován) pacienta s oteklým nádorem (označeným jako e) na pravé hlasivce. Odraz světla označující vrchol slizniční vlny (rm) je zde daleko od hranice hlasivkové oblasti (oboustranná šipka). To značí, že vlna je na hraně omezoována přítomností otoku. Obrázek a popis obrázku převzat z [6].



Obrázek 3.30: Typ slizniční vlny levé hlasivky: slizniční vlna bez světelných odrazů, slizniční vlna se světelnými odrazy, slizniční vlna není přítomna. Převzato z [4].



Obrázek 3.31: Kategorie dosahu slizniční vlny levé hlasivky bez světelných odrazů (vlevo) a kategorie dosahu slizniční vlny levé hlasivky se světelnými odrazy (vpravo). Převzato z [4].



Obrázek 3.32: Kategorie vzdálenosti slizniční vlny levé hlasivky se světelnými odrazy od amplitudy otevření hlasivkové štěrbiny: slizniční vlna začíná od kraje středového vrcholu, je blízko od středového vrcholu, je daleko od středového vrcholu a začíná za středovým vrcholem - tj. je zpožděná. Převzato z [4].

8. Poměr délky trvání otevírání hlasivek a délky trvání zavírání hlasivek

Dle [5] by tento poměr mohl odpovídat tlaku hlasivek v například v závislosti na svalové slabosti. Důležité případy jsou zobrazeny na obrázku 3.33. Na obrázku

3.34 jsou snímky hlasivek s rozdílným poměrem otevírání a zavírání.



Obrázek 3.33: Kategorie poměru délky otevírání a zavírání levé hlasivky: otevírání je značně kratší, kratší, mírně kratší, stejné jako zavírání. Převzato z [4].

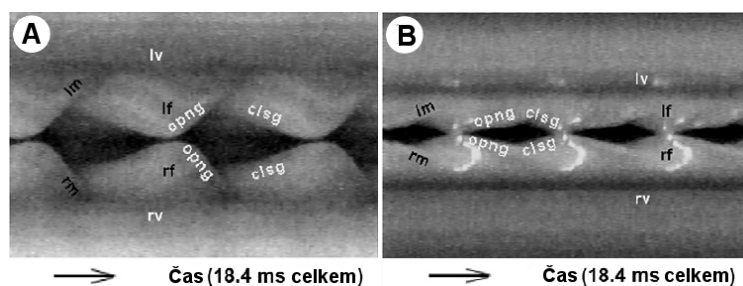
9. Tvar středových vrcholů

Středové vrcholy jsou přítomny pouze pokud se hlasivky nezavírají. Podobně jako u bočních vrcholů mohou být ostré nebo zakulacené 3.35. Snímky hlasivek s ostrými a neostrými středovými vrcholy jsou zobrazeny na obrázku 3.36. Poruchy tvaru středových vrcholů jsou často spojeny s poruchou tvaru středového povrchu hlasivek jako je dvojitý středový vrchol a středová nehladkost.

10. Poruchy průběhu vibrace které nemusí způsobovat poruchy v periodicitě vibrací

Tyto poruchy způsobují změny v jinak jednoduchém tvaru vibrací. Patří mezi ně vlnění průběhu vibrace, dvojitý středový vrchol a středová nehladkost 3.37.

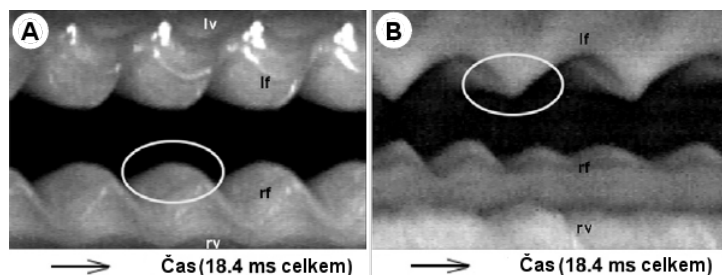
Výskyt těchto poruch je brán jako varovný znak zhoršené funkce hlasivek. Jsou zobrazeny na obrázku 3.38.



Obrázek 3.34: Rozdíl mezi dobou otevírání (opng) a zavírání (clsg) hlasivky. **A)** Hlasivky muže s omezenou pohyblivostí levé hlasivkové chrupavky. **B)** Hlasivky ženy s normálně vypadajícím hrtanem, která si stěžovala na unavenost hlasu. Obrázek a popis obrázku převzat z [6].



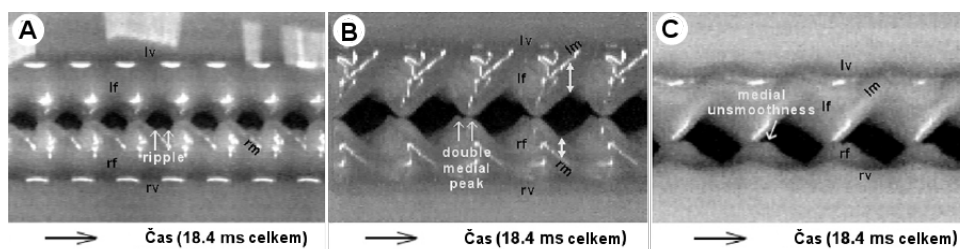
Obrázek 3.35: Kategorie tvaru středových vrcholů: zakulacený, spíše zakulacený, spíše ostrý a ostrý středový vrchol levé hlasivky. Převzato z [4].



Obrázek 3.36: Tvar středových vrcholů (zakroužkované). **A)** Zakulacený vrchol hlasivek ženy s normální strukturou hlasivek a tradiční diagnózou psychogenní ztráty hlasu. **B)** Ostrý vrchol hlasivek [dysphonic] muže s hlasivkami s defektivní strukturou. Obrázek a popis obrázku převzat z [6].



Obrázek 3.37: Poruchy průběhu vibrací levé hlasivky: vlnění (dvojitý boční vrchol), dvojitý středový vrchol a středová nehladkost. Převzato z [4].



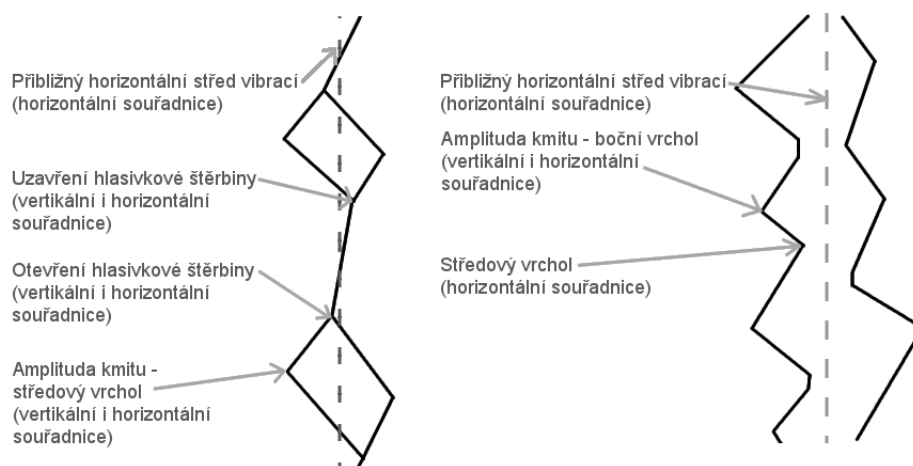
Obrázek 3.38: Poruchy průběhu vibrace, které nemusí způsobovat poruchy v periodicitě vibrací. **A)** Vlnění průběhu vibrace ("ripple"), dvojitý středový vrchol (**B**) a středová nehladkost (**C**) pacientů s nádorem hlasivky. Obrázek a popis obrázku převzat z [6].

3.4 Cíle detekce

Největší přínos má automatická detekce parametrů, které je pro jejich přesné určení nutno ve snímku naměřit a jejichž nepřesný odhad by mohl způsobit špatné vyhodnocení videokymografického snímku. Naopak detekce parametrů, které je možné odhadnout nebo rozeznat podle tvaru, není tak důležitá. V tomto oddílu jsou vybrány parametry z dotazníku pro vyhodnocování videokymografických snímků vhodné pro automatickou detekci.

Navržená automatická detekce vychází z určení vybraných příznaků ve videokymografickém snímku, takzvaných *základních příznaků*. Jsou to příznaky, jejichž poloha není přímo odvoditelná ze znalosti ostatních příznaků. Základní příznaky jsou znázorněny na obrázku 3.39. Jsou to:

- přibližný horizontální střed vibrací
- horizontální a vertikální souřadnice otevření a uzavření hlasivkové štěrbiny
- pro každou hlasivku vertikálních polohy středových vrcholů
- horizontální a vertikální souřadnice amplitud kmitů



Obrázek 3.39: Základní detekované příznaky na snímku, kde dochází k uzavření hlasivkové štěrbiny (vlevo) a na snímku, kde nedochází k uzavření hlasivkové štěrbiny (vpravo).

Ze znalosti těchto základních příznaků je možné odvodit hodnoty dalších charakteristik vibrací:

- přítomnost vibrací a počet vibrací každé hlasivky
- variabilita průběhu po sobě jdoucích vibrací
- poměr délky otevření k délce uzavření hlasivky
- přítomnost levo-pravé asymetrie

- rozdíl amplitudy levé a pravé hlasivky
- rozdíl frekvence levé a pravé hlasivky
- fázový posun kmitání levé a pravé hlasivky
- poměr délky otevírání hlasivky k délce zavírání hlasivky
- posun osy

Cílem práce je především extrahovat základní příznaky a z nich určit odvozené charakteristiky vibrací. Je přitom velmi důležité, aby extrakce základních příznaků byla dostatečně přesná. Extrakce některých základních příznaků se opírá o znalost jiných. Nepřesná extrakce jednoho základního příznaku tak může vést jednak k chybné extrakci jiných základních příznaků a jednak k chybnému určení odvozených charakteristik vibrací. Pokud je například chybně určen střed vibrací, jsou nutně špatně extrahovány všechny další základní příznaky. Z tohoto pohledu je nezbytná možnost manuálního zásahu do průběhu algoritmu. Polohu základních příznaků by mělo být možné manuálně opravit.

3.5 Vyhodnocování výsledků

Protože se základní příznaky navzájem ovlivňují, je nutné v případě změny algoritmu detekce daného základního příznaku nejen vyhodnotit správnost detekce tohoto příznaku, ale také vliv změněné detekce příznaku na detekci dalších příznaků. Hlasivky různých pacientů mají navíc často velmi různý průběh vibrací a snímky jsou tak vzájemně značně odlišné. Algoritmy, které dobře fungují na jednom snímku mohou zcela selhávat na ostatních a je tedy opět nutné vyhodnotit správnost detekce parametrů na dostatečně reprezentativním souboru snímků.

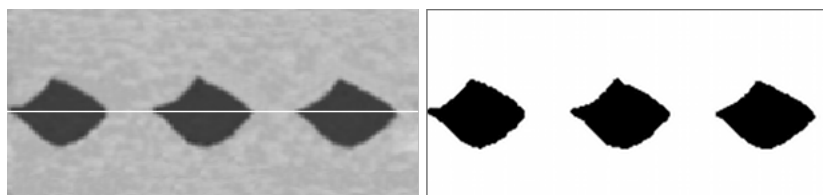
Proto byl vytvořen systém, který umožňuje vyhodnotit správnost automatické extrakce. Systém vyhodnotí proběhlou detekci na základě porovnání automaticky určených a známých hodnot příznaků. Tak je možné vyvinout dostatečně robustní a tím přesnější algoritmy.

Kapitola 4

Popis řešení

4.1 Možnosti extrakce základních příznaků

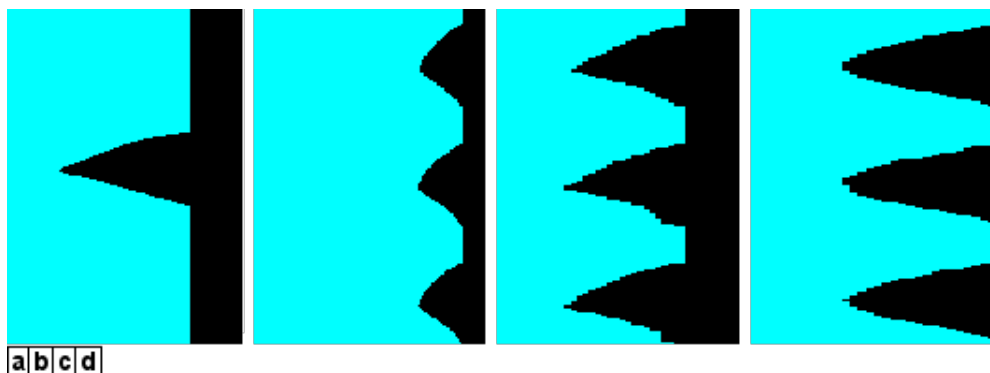
V oddílu 3.4 byly stanoveny základní příznaky, které je třeba extrahovat. Díky tomu, že videokymografický snímek se skládá ze dvou ohraničených oblastí - hlasivkové štěrby a jejího okolí, zdá se vhodné nejprve tyto dvě oblasti segmentovat a následně určit všechny základní příznaky hledáním hodnot - typicky lokálních minim nebo lokálních maxim - ve vektorech řádkových a sloupcových součtů segmentovaného snímku. Na obrázku 4.1 je zobrazen “ideální” videokymografický snímek a obrázek vzniklý jeho segmentací. Na obrázku 4.2 jsou sloupcové a řádkové součty segmentovaného snímku a sloupcové součty horní a spodní poloviny videokymografického snímku. Snímek je rozdělen v místě minima řádkových součtů.



Obrázek 4.1: Ideální snímek hlasivek s vyznačeným středem vibrací (vlevo), segmentovaný ideální snímek (vpravo).

Segmentace reálného videokymografického snímku je bohužel problematická. Snímek je velmi zašuměný, je na něm přítomno množství artefaktů způsobených například interferencí okolí s hlasivkami a hranice hlasivkové štěrby je na některých místech neostrá. Okolní tkáň není homogenní a na některých místech není světlejší než hlasivková štěrbina. Tato místa tvoří na videokymografickém snímku tmavé vertikální pruhy. Snímky navíc nebývají přesně vystředěné a hlasivková štěrbina se tak nachází mimo střed snímku. Další problémy vzniklé při použití technik segmentace na videokymografický snímek byly popsány v oddílu 3.2.

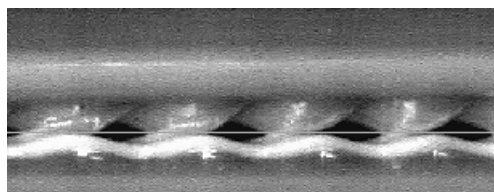
Alternativní přístup k segmentaci videokymografického snímku je použití řádkových a sloupcových součtů původního, nesegmentovaného snímku a určení základních příznaků z těchto součtů. Na obrázku 4.3 je reálný snímek hlasivek a



Obrázek 4.2: Řádkové součty ideálního snímku hlasivek (a), sloupcové součty ideálního snímku hlasivek (b), sloupcové součty horní poloviny ideálního snímku hlasivek (c) a sloupcové součty spodní poloviny ideálního snímku hlasivek (d).

na obrázku 4.4 jsou zobrazeny jeho řádkové a sloupcové součty a sloupcové součty horní a spodní poloviny snímku. I když snímek obsahuje velké množství artefaktů a jeho přesná segmentace by byla problematická, zachovávají řádkové a sloupcové součty z velké části stejný základní tvar jako součty segmentovaného snímku. Při použití řádkových či sloupcových součtů přímo na původní snímek dochází k přirozené eliminaci vlivu šumu a ve srovnání s prahováním je tento přístup méně citlivý na rozmazání okrajů hlasivkové štěrby snímku. Nabízí se tedy možnost extrahovat základní příznaky přímo z těchto součtů, bez použití segmentace. Tento přístup je ovšem také problematický, protože zejména vlivem artefaktů ve snímku dochází ke zkreslení součtů. Například v případě řádkových součtů tak může být přítomno větší množství lokálních minim a střed snímku nelze určit jako minimum řádkových součtů.

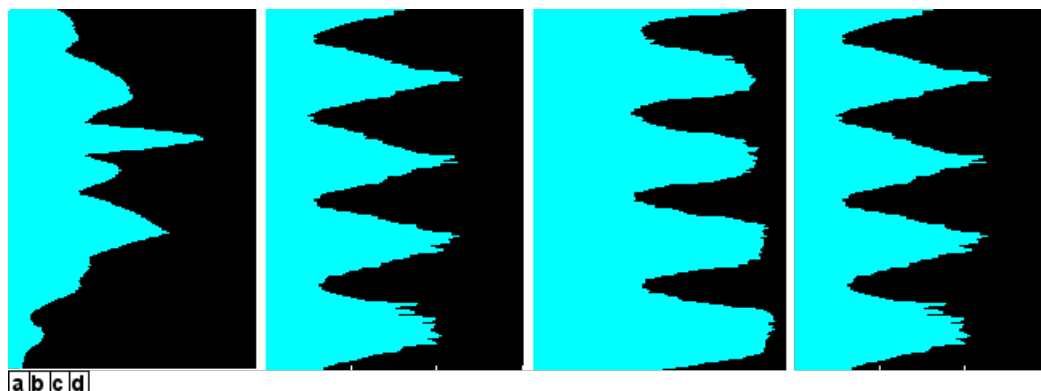
Vhodné je oba přístupy kombinovat. Navrhovaný postup vychází z postupného určování základních příznaků. Jednotlivé příznaky postupně rozdělují snímek na menší části a další základní příznaky se tak hledají pouze v daných výřezech snímku. V rámci daného výřezu se použije buď přístup založený na řádkových a sloupcových součtech nebo vhodná technika segmentace.



Obrázek 4.3: Reálný snímek hlasivek s vyznačeným středem vibrací.

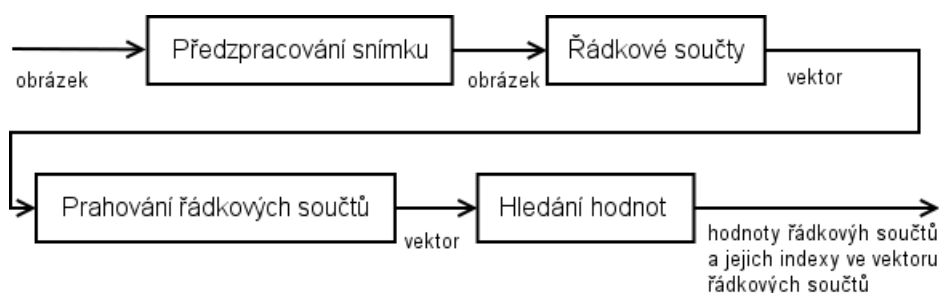
4.2 Základní algoritmus a jeho konfigurace

V každém dílčím kroku postupu určování základních příznaků se použije vhodně modifikovaná kombinace segmentace a přístupu založeném na řádkových či sloup-



Obrázek 4.4: Řádkové součty reálného snímku hlasivek (a), sloupcové součty reálného snímku hlasivek (b), sloupcové součty horní poloviny reálného snímku hlasivek (c) a sloupcové součty spodní poloviny reálného snímku hlasivek (d).

cových součtech. Byl vytvořen univerzální algoritmus, který je možné nakonfigurovat tak, aby danou kombinaci těchto přístupů provedl. Algoritmus také umožňuje jednoduchou vizualizaci průběhu. Schéma algoritmu je na obrázku 4.5.



Obrázek 4.5: Algoritmus řádkových součtů

Konfigurace algoritmu je realizována pomocí hierarchie konfiguračních objektů. Každá volba má svoje výchozí hodnoty a při novém použití algoritmu tak není nutné konfigurovat všechny aspekty algoritmu. Jednotlivé konfigurace lze měnit nezávisle na sobě a lze tak snadno experimentovat s různými kombinacemi metod. Konfigurace jednotlivých použití algoritmu je při dávkové detekci hodnot základních parametrů - stejně jako detekované hodnoty parametrů - zapsána do textového souboru. Je tedy možné zpětně dohledat, jaké konfigurace algoritmu vedly k daným výsledkům. Dále je uveden podrobnější popis fází algoritmu:

1. Předzpracování snímku

Dále obsahuje tyto samostatně konfigurovatelné fáze jejichž pořadí lze libovolně měnit:

- potlačení šumu
- segmentace obrázku

- **sémantické zpracování obrázku**
- **invertování obrázku**

Ve fázi sémantického zpracování lze snímek předzpracovat tak, aby řádek, který je tvořen delší souvislou řadou tmavých pixelů začínající vlevo měl vyšší hodnotu součtu pixelů na tomto řádku, než řádek s kratší cestou nebo tak, aby byl zvýrazněn vliv pixelů určitých intenzit. Invertování obrázku slouží typicky k přepínání mezi hledáním lokálních minim a lokálních maxim.

2. Spočtení řádkových součtů obrázku

3. Prahování řádkových součtů

4. Hledání hodnot ve vektoru řádkových součtů

Algoritmus z výsledného vektoru řádkových součtů vybere hodnoty a jejich indexy. Pro výběr hodnot z vektoru řádkových součtů lze zvolit jeden z těchto algoritmů:

- **algoritmus hledání lokálních minim**
- **algoritmus hledání nulových úseků**

Algoritmus hledání lokálních minim lze dále konfigurovat. Lze například povolit ověřování nalezeného lokálního minima. To je založené na porovnání průměrů pixelů v okolí nalezeného minima a okolí pixelů nalevo a napravo od nalezeného minima. Pokud nalezené lokální minimum nesplňuje dané požadavky, hledá algoritmus další lokální minimum. Algoritmus hledání nulových úseků hledá ve vektoru souvislé úseky složené z nul. Jeho konfigurace zahrnuje například minimální délku úseku.

4.3 Hledání prahu pro segmentaci hlasivkové štěrbiny

Pro segmentaci založenou na prahování je nutné stanovit hladinu prahu. V případě segmentace hlasivkové štěrbiny od jejího okolí je to intenzita pixelů na přechodu mezi těmito dvěma částmi. Metody hledání prahu v oddílu 3.2.3 předpokládají, že snímek se skládá ze dvou oblastí s relativně homogenní intenzitou a v histogramu jsou tak patrné dva dostatečně od sebe oddělené vrcholy. V případě videokymografických snímků ale tento předpoklad není splněn - okolí hlasivky netvoří homogenní oblast a histogram se tak skládá z více vrcholů. Navržený algoritmus vychází z následujících vlastností videokymografického snímku:

- Oblast hlasivkové štěrbiny je nejtmavší oblast ve snímku. V histogramu bude tedy tvořit nejlevější vrchol. Tato oblast je relativně homogenní, proto by tento vrchol měl být dobře patrný. Na druhou stranu může hlasivková

štěrbina v některých případech zabírat menší část výřezu snímku než okolí hlasivky a nelze tedy předpokládat, že tento vrchol bude největší vrchol histogramu.

- Na přechodu mezi hlasivkovou štěrbinou a jejím okolím dochází k změně intenzit pixelů. V této oblasti je tedy relativně malé množství pixelů s podobnou intenzitou a vrchol histogramu odpovídající hlasivkové štěrbíně je separován od ostatních vrcholů údolím.

Algoritmus pracuje v těchto fázích:

1. Vytvoření histogramu výřezu

Vytvoří se histogram výřezu s relativně velkým počtem přihrádek. Osvědčila se hodnota 128.

2. Vyhlazení histogramu

Hodnota histogramu po vyhlazení je v každé v přihrádce rovna maximu z hodnot v okolí přihrádky. V algoritmu je použito okolí o velikosti 4. Vyhlazení histogramu přispěje ke zvýšení robustnosti celého postupu.

3. Hledání vrcholu histogramu odpovídajícímu hlasivkové štěrbíně

Je nalezena hodnota prvního lokálního maxima ve vyhlazeném histogramu. Ta za platnosti výše uvedených pozorování odpovídá vrcholu histogramu tvořeným pixely hlasivkové štěrbiny.

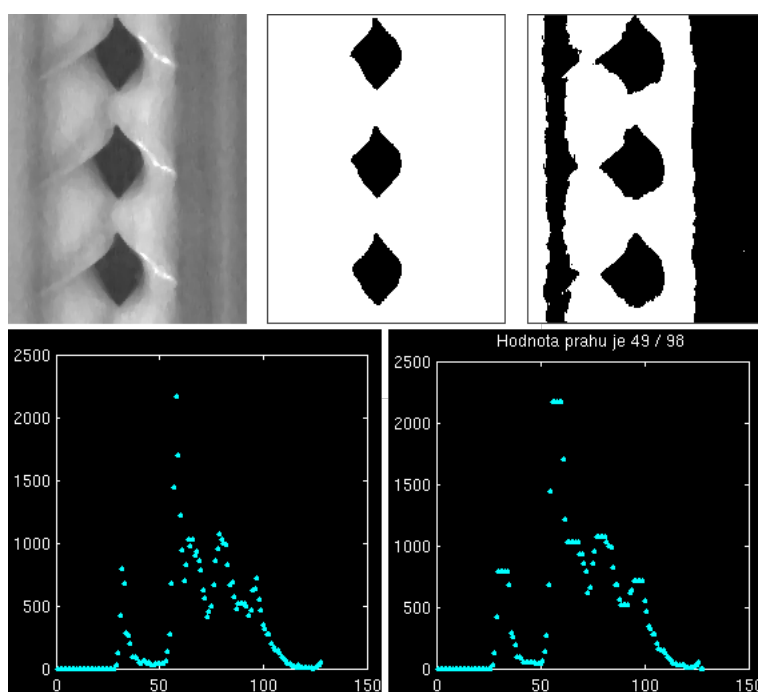
4. Hledání intenzity odpovídající přechodu mezi hlasivkovou štěrbínou

Podle pozorování výše je hodnota intenzity oddělující hlasivkovou štěrbinu od jejího okolí v obrázku málo zastoupená a odpovídající přihrádka histogramu tedy tvoří lokální minimum napravo od vrcholu odpovídajícímu hlasivkové štěrbíně. Algoritmus tedy hledá ve vyhlazeném histogramu první lokální minimum ležící napravo od tohoto vrcholu.

Při vytváření histogramu výřezu se použije velký počet přihrádek proto, že oblast hlasivkové štěrbiny je relativně homogenní a pixely do ní patřící v takovém histogramu vytvoří výrazný vrchol. V případě výrazné nehomogenity oblasti hlasivkové štěrbiny může dojít k nalezení prahu, který při následné segmentaci segmentuje pouze část hlasivkové štěrbiny. Nejhorší situace nastane, pokud oblast okolí hlasivkové štěrbiny obsahuje pixely stejné intenzity jako oblast přechodu

mezi hlasivkovou štěrbinou a jejím okolím. To se může stát ve vzdálenějších částech okolí hlasivkové štěrbiny. Údolí za vrcholem histogramu odpovídajícím hlasivkové štěrbině pak nemusí být patrné. Výřez snímku z kterého se počítá práh by proto neměl tyto oblasti obsahovat. V případě, že výše uvedené předpoklady nejsou splněny a nalezený práh není správný, použije se pro hledání prahu Otsuova metoda 3.2.3 a následně se vybere správnější z určených prahů. Správnost prahu algoritmus určí podle podílu pixelů přiřazených do segmentu příslušející hlasivkové štěrbině. V konfiguraci lze pro každý případ hledání prahu nastavit minimální a maximální hodnotu tohoto podílu pixelů.

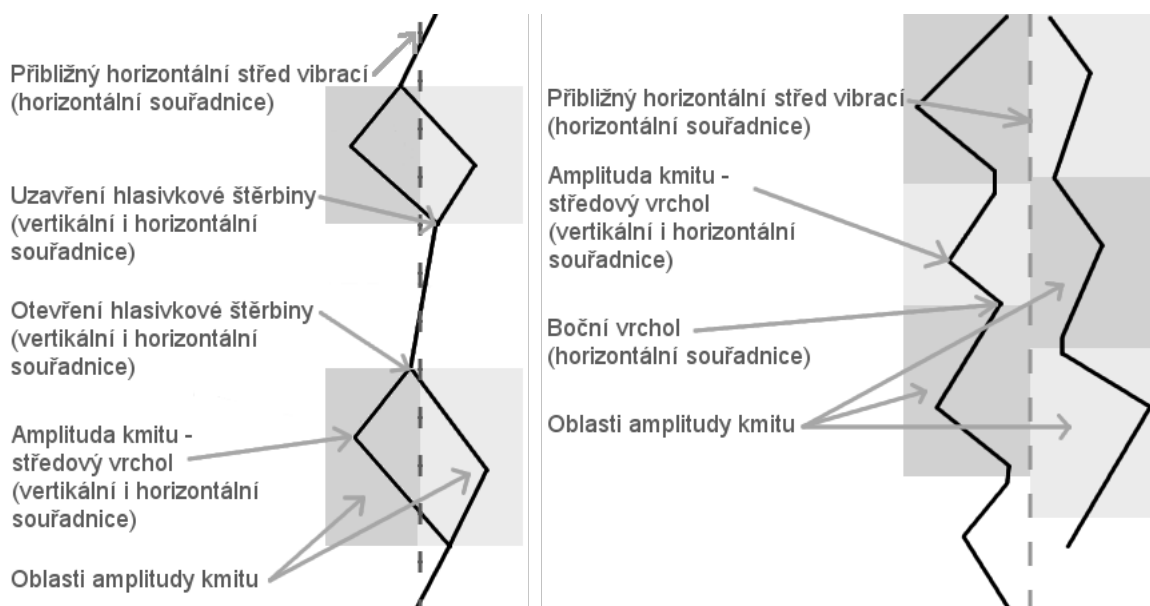
Na obrázku 4.6 je histogram snímku, vyhlazený histogram snímku a segmenty vzniklé prahováním prahem určeným pomocí navrženého algoritmu a podle Otsuovy metody. Histogram snímku obsahuje více vrcholů histogramu, nesplňuje tak předpoklady Otsuovy metody a práh určený dle navrženého algoritmu je tak lepší.



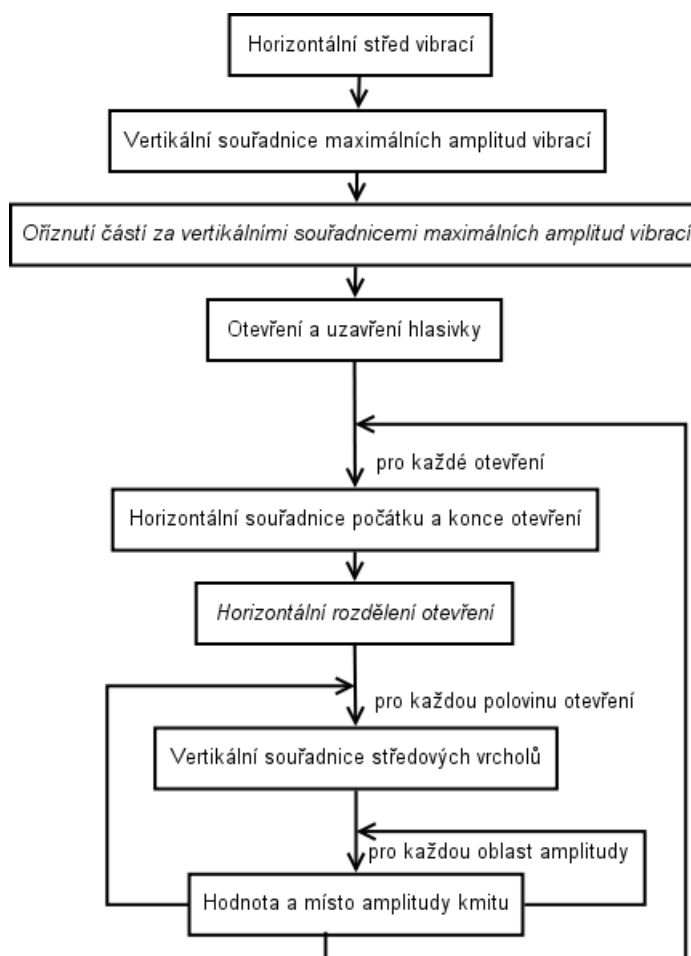
Obrázek 4.6: Snímek hlasivek s potlačeným šumem, snímek prahovaný pomocí prahu získaného navrženým algoritmem hledání prahu pro segmentaci hlasivkové štěrbiny, snímek prahovaný pomocí prahu získaného pomocí Otsuovy metody, histogram snímku se 128 přihrádkami, vyhlazený histogram.

4.4 Extrakce základních příznaků ze snímku

Všechny základní příznaky jsou označeny na schematicém obrázku videokymografického snímku hlasivek 4.7. Průběh celého algoritmu extrakce základních příznaků ze snímku je znázorněn v diagramu 4.8.



Obrázek 4.7: Základní příznaky extrahované z videokymografického snímku.



Obrázek 4.8: Schéma algoritmu extrakce základních příznaků ze snímku.

V průběhu extrakce dojde nejprve k nalezení přibližného horizontálního středu vibrací. Dále se naleznou vertikální souřadnice maximální amplitudy vibrací levé i pravé hlasivky. Ty vymezují oblast ve snímku, kde se nacházejí všechny další základní příznaky. Okolní oblasti jsou tedy oříznuty.

Následně jsou hledána místa otevření a uzavření hlasivky. Pro každé otevření se naleznou horizontální souřadnice začátku i konce. Otevření se podle těchto souřadnic rozdělí na část příslušející levé hlasivce a část příslušející pravé hlasivce. Část odpovídající levé hlasivce se horizontálně překlopí. Je pak stejně orientovaná jako část odpovídající pravé hlasivce a při dalším zpracování není třeba rozlišovat mezi různými orientacemi. Protože levá hlasivka může kmitat zcela jinak než pravá, pracuje se dále s částmi příslušejícími levé a pravé hlasivce samostatně.

V každé části každého otevření jsou nalezeny vertikální souřadnice středových vrcholů. Středové vrcholy vymezují takzvané *místa amplitudy* - části snímku ve kterých se nachází jedna amplituda kmitu hlasivku. Pokud není v části otevření příslušející dané hlasivce žádný středový vrchol, je v celém otevření jedna amplituda kmitu. V rámci každého místa amplitudy se nalezne hodnota amplitudy a horizontální souřadnice amplitudy.

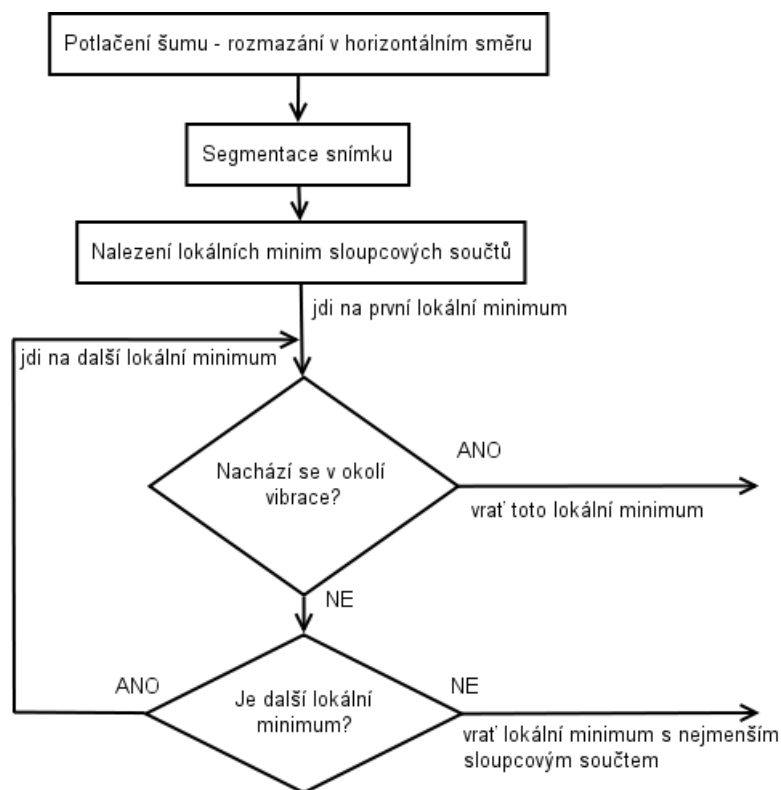
Hledání středu vibrací



Schéma algoritmu je znázorněno na diagramu 4.9. Algoritmus pracuje v rámci celého videokymografického snímku. Nejprve se potlačí šum pomocí mediánového filtru. Použije se obdélníkového okna s větším počtem řádků než sloupců. Tím dojde k rozmazání snímku ve vertikálním směru což vzhledem ke tvaru snímku a tomu, že algoritmus hledá horizontální souřadnici, nevadí.

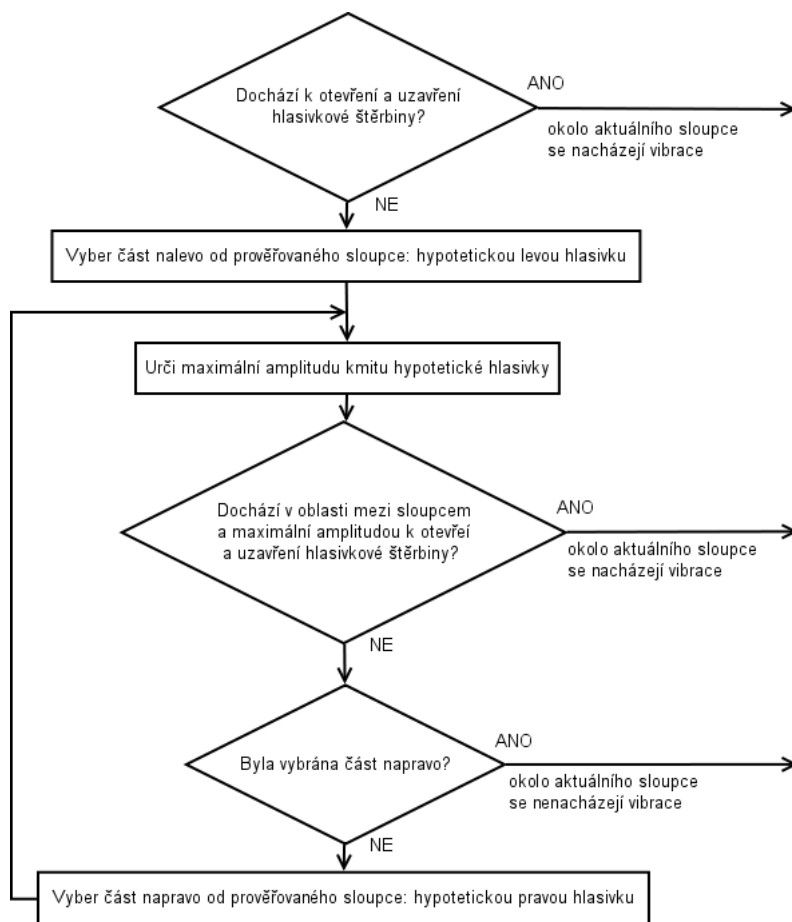
Další krok je segmentace snímku pomocí prahování. Použitý práh je spočítán pomocí algoritmu 4.3. V takto oprahovaném snímku je dále nalezeno několik lokálních minim sloupcových součtů snímku. Tato minima se prochází v pořadí od minima s nejnižší hodnotou řádkového součtu a vybere se z nich to minimum, pro které platí, že se okolo odpovídajícího sloupce nacházejí vibrace. Pokud mezi nalezenými minimy žádné takové není, vybere se minimum s nejnižší hodnotou řádkového součtu.

To, jestli se okolo daného sloupce nachází vibrace, se určí na základě algoritmu znázorněného na obrázku diagramu 4.10. Algoritmus se nejprve pokusí vyhledat hypotetická místa uzavření a otevření hlasivkové štěrby v daném sloupci. Pokud je hledání neúspěšné, stále může být dané místo středem vibrací. V místě středu vibrací totiž k otevření a uzavření hlasivkové štěrby nemusí docházet. Proto se dále zjišťuje, zda kmitá hypotetická levá hlasivka. K tomu se nejprve určí maximální amplituda kmitu této hypotetické hlasivky. Dále dojde k hledání míst otevření a uzavření hlasivky v místě mezi hypotetickým středem hlasivky a místem hypotetické maximální amplitudy kmitu. Pokud je hledání kmitů na levé straně neúspěšné, stejným způsobem se hledají kmity pravé hlasivky. Pokud nejsou nalezeny ani kmity pravé hlasivky, okolo daného sloupce se vibrace nenacházejí.



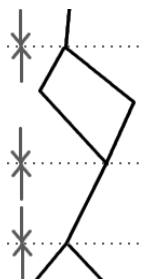
Obrázek 4.9: Schéma algoritmu extrakce středu snímku.

Hledání míst otevření a uzavření a hledání místa maximální amplitudy probíhá podle algoritmů popsaných dále.

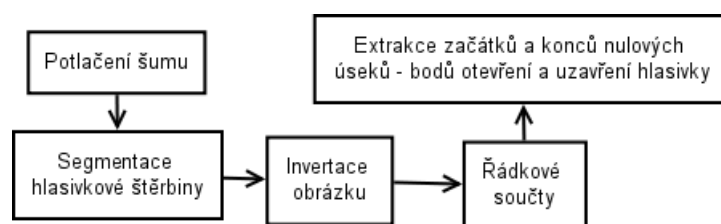


Obrázek 4.10: Schéma algoritmu určení, zda jsou okolo daného sloupce přítomny vibrace.

Hledání vertikálních souřadnic otevření a uzavření hlasivek



Nejprve dojde k potlačení šumu pomocí adaptivního mediánového filtru. Použije se obdélníkové okno s větším počtem řádků. Dojde tak k vyhlazení šumu a zachování hran otevření a uzavření. Dále dojde k segmentaci hlasivkové štěrby pomocí algoritmu narůstání oblastí 3.2.4. Pro nalezení semínek se použije informace o poloze středu snímku a dále práh určený za použití algoritmu 4.3. Semínka jsou tvořena pixely ležící v blízkosti středu s intenzitou menší než nalezený práh.



Obrázek 4.11: Schéma algoritmu hledání míst otevření a uzavření hlasivek.

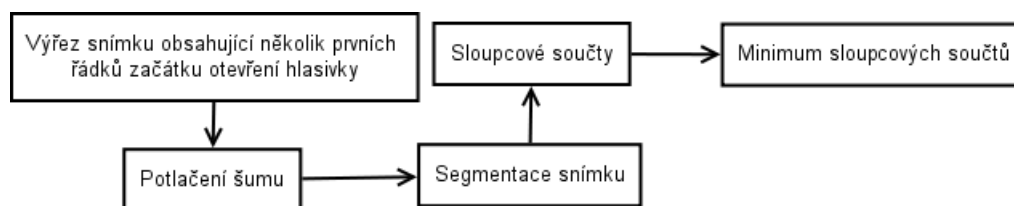
Následně se spočítají řádkové součty segmentovaného snímku. Prvky vektoru řádkových součtů s hodnotami nula odpovídají řádkům na kterých je hlasivková štěrbina zavřená, prvky s hodnotami vyššími než nula odpovídají řádkům na kterých je hlasivková štěrbina otevřená. Z vektoru se extrahují začátky a konce úseků otevření hlasivek a na jejich základě se určí místa otevření a uzavření hlasivek.

Hledání horizontálních souřadnic otevření a uzavření hlasivek



Vytvoří se výřez snímku obsahující několik prvních řádků začátku otevření hlasivky, v tomto výřezu se potlačí šum pomocí adaptivního mediánového filtru. Použije se okno obdélníkového tvaru s větším počtem řádků než sloupců. K většímu rozmazání tedy dojde ve vertikálním směru. Snímek se snímek segmentuje metodou narůstání oblastí. Využije se přitom informace, že hlasivková štěrbina se nachází ve spodní části výřezu v případě otevření hlasivky a horní části výřezu v případě uzavření hlasivky. Dále se nalezne minimum sloupcových součtů v tomto výřezu. Hledání horizontálních souřadnic konce otevření hlasivky probíhá analogicky.

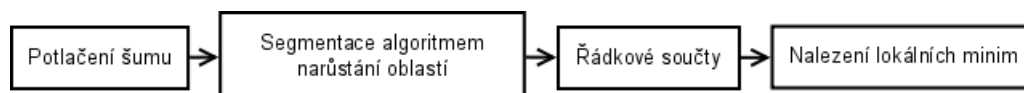
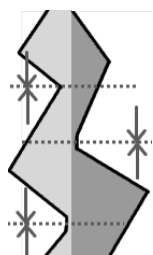
Po nalezení těchto souřadnic se otevření hlasivky rozdělí na levou a pravou část. Hledání středových vrcholů a amplitud pak probíhá v obou těchto částech odděleně. Aby nebylo nutné rozlišovat mezi levou a pravou částí, dojde navíc k překlopení levé části. Algoritmy pak mohou na levé části otevření pracovat stejně jako na pravé části.



Obrázek 4.12: Schéma algoritmu hledání horizontálních souřadnic otevření hlasivky.

Hledání míst středových vrcholů

Nejprve dojde k potlačení šumu ve výřezu pomocí adaptivního mediánového filtru. Použije se obdélníkové okno s větším počtem sloupců. Dojde tak k vyhlazení šumu a zachování tvaru středových vrcholů.



Obrázek 4.13: Schéma algoritmu hledání míst středových vrcholů.

Dále je oblast hlasivkové štěrbině segmentována pomocí algoritmu narůstání oblastí 3.2.4. Využije se přitom informace, že nejlevější sloupec výřezu odpovídá středu otevření hlasivky. Jako semínka jsou proto zvoleny pixely ležící u levého okraje výřezu s intenzitou menší než práh vypočtený algoritmem 4.3. Nakonec dojde k nalezení lokálních minim řádkových součtů prahovaného výřezu mající od sebe danou minimální vzdálenost. Pro hledání míst středových vrcholů je povoleno ověřování lokálních minim a nakonfigurováno tak, aby byla vybrána pouze lokální minima skutečně odpovídající středovým vrcholům.

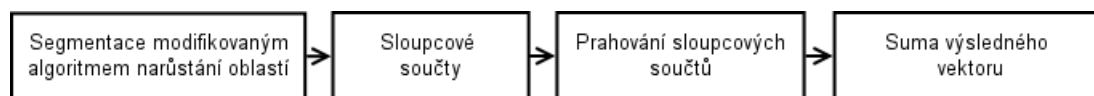
Hledání hodnoty amplitudy

Algoritmus pracuje v rámci takzvaných míst amplitud - výřezů snímku na kterých je právě jedna amplituda kmitu. Výřez se nejprve segmentuje za použití modifikovaného algoritmu narůstání oblastí 3.2.4. Narůstání oblastí využívá stejně jako v případě hledání vertikálních souřadnic informaci, že hlasivková štěrbina začíná na levé straně výřezu. Jako semínka jsou tedy zvoleny pixely na levé straně výřezu jejichž intenzita je menší než daný práh. Hodnota prahu se spočítá za použití algoritmu 4.3.



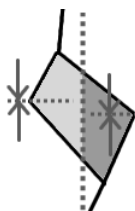
Dále se spočítají sloupcové součty výřezu. Na vzniklý vektor sloupcových součtů se aplikuje algoritmus prahování s prahem blízkým nule. Tím dojde k redukci chyby vzniklé zubatými okraji v segmentovaném snímku zapříčiněnými neostrotí přechodu mezi hlasivkovou štěrbinou a jejím okolím. Protože má výsledný vektor jedničky na místech odpovídajícím sloupcům výřezu na kterých je část hlasivkové

šterbiny a nuly na místech odpovídajícím sloupcům na kterých je pouze okolí hlasivkové šterbiny, je hodnota amplitudy daná součtem prvků tohoto vektoru.



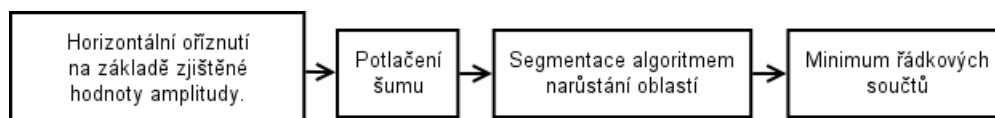
Obrázek 4.14: Schéma algoritmu hledání hodnoty amplitudy.

Hledání vertikální polohy amplitudy



Nejprve se místo amplitudy na základě zjištěné hodnoty amplitudy horizontálně ořízne. Ve vzniklém výřezu se potlačí šum za použití adaptivního mediánového filtru. Použije se okno obdélníkového tvaru s větším počtem sloupců než řádků. K většímu rozmazání tedy dojde v horizontálním směru.

Dále dojde k segmentaci hlasivkové šterbiny za použití algoritmu narůstání oblastí 3.2.4. Při hledání semínek se stejně jako v případě hledání hodnoty amplitudy se využije informace, že hlasivková šterbina se nachází na levé straně výřezu. Vertikální souřadnice amplitudy se pak určí jako minimum řádkových součtů segmentovaného výřezu.



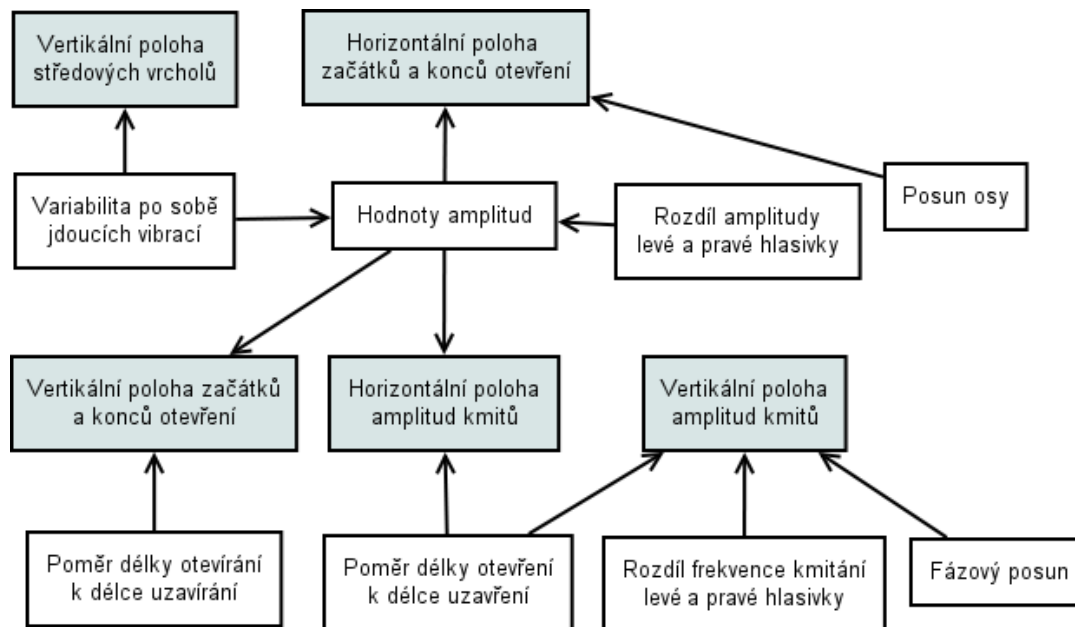
Obrázek 4.15: Schéma algoritmu hledání vertikální polohy amplitudy.

Hledání maximální amplitudy hlasivky

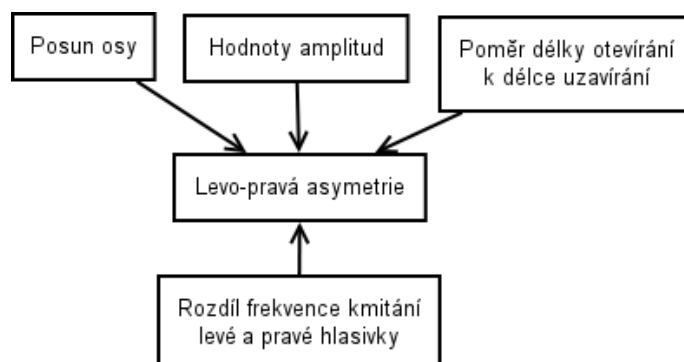
Hledání maximální amplitudy hlasivky probíhá analogicky jako hledání hodnoty amplitudy. Místo použití sloupcových součtů přes řádky oblasti amplitudy se ale použijí sloupcové součty přes všechny řádky ve snímku.

4.5 Odvozené charakteristiky kmitání

Z extrahovaných základních příznaků lze získat řadu odvozených charakteristik kmitání. Extrakce těchto charakteristik je znázorněna na obrázcích 4.16 a 4.17. Šipky na obrázcích směřují z příznaku nebo odvozené charakteristiky kmitání použité pro extrakci do extrahované odvozené charakteristiky.



Obrázek 4.16: Extrakce odvozených charakteristik kmitání.



Obrázek 4.17: Detekce Levo-pravé asymetrie.

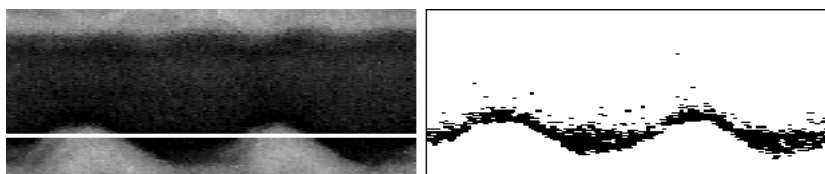
Kapitola 5

Vyhodnocení

Za účelem vyhodnocení implementované extrakce charakteristik kmitání ve videokymografickém snímku byl shromážděn soubor testovacích snímků a na těchto snímcích byly naměřeny hodnoty základních příznaků. Tyto hodnoty byly následně porovnány s hodnotami získanými automatickou extrakcí.

Nejdůležitější základní příznak je střed vibrací. Pokud je střed vibrací určen chybně, jsou chybně určeny i všechny další charakteristiky. Automatická extrakce středu vibrací byla proto podrobně testována na největším souboru snímků. Z 89 videokymografických snímků byla poloha středu vibrací chybně určena pouze na jednom snímku. Tento snímek je zobrazen na obrázku 5.1-vlevo. K chybě došlo ve fázi segmentace snímku. Nalezený segment hlasivkové štěrbiny je zobrazen na obrázku 5.1-vpravo. Vinou nehomogenní intenzity pixelů oblasti hlasivkové štěrbiny je segmentována pouze spodní, tmavší část této štěrbiny a následně je střed extrahován z této části.

Chyba v extrakci středu vibrací by mohla teoreticky nastat i pokud nedochází k uzavření hlasivkové štěrbiny a hlasivková štěrbina je i ve svém nejužším místě široká. I v případě, že hlasivková štěrbina by byla správně segmentována, mohl by být vybrán řádek blízko jedné z kmitajících hlasivek. Řádky v jejichž okolí se nacházejí vibrace jsou totiž při hledání středu vibrací preferovány. Na testovaném souboru se tato chyba ale neprojevila.



Obrázek 5.1: Extrakce středu snímku. Snímek s vyznačenou polohou nalezeného středu (vlevo), chybně segmentovaná hlasivková štěrbina (vpravo).

V tabulce 5.1 jsou statistiky jednotlivé extrakce ostatních základních příznaků. Při jednotlivé extrakci byly vždy extrahovány pouze hodnoty jednoho příznaku a hodnoty ostatních příznaků byly převzaty z manuálního měření. V těchto statistikách tak nedochází k chybnému určení jednoho příznaku z důvodu chybného určení jiného příznaku. V tabulce 5.2 jsou statistiky plně automatické extrakce

všech základních příznaků. Zde již k chybnému určení jednoho příznaku z důvodu chybného určení jiného dojít může a úspěšnost plně automatické extrakce je tak u většiny příznaků nižší. Rozdíl většinou ale není nijak markantní. To svědčí o dobré robustnosti nástroje. Rozdíl je více patrný u příznaků při jejichž určování je větší vliv správného určení jiných příznaků. Při určování vertikální souřadnice začátku a konce otevření je dokonce lepší úspěšnost při plně automatické detekci. To ukazuje na to, že automaticky nalezený střed je z hlediska následného určení souřadnic začátku a konce otevření lepší než manuálně určený.

Extrakce všech zmíněných příznaků byla prováděna na stejném souboru 50 videokymografických snímků. Ve sloupci Počet obou tabulek je počet snímků na kterých se vyskytovala alespoň jedna hodnota daného příznaku a celkový počet hodnot tohoto příznaku na všech snímcích. Ve sloupci Úspěšnost je podíl snímků bez chybně detekované hodnoty z celkového počtu snímků, podíl všech správně detekovaných hodnot z celkového počtu hodnot a podíl všech správně detekovaných hodnot na snímcích, kde byl detekovaný správný počet hodnot daného příznaku. Ve sloupci Průměrná odchylka je průměrná odchylka správně extrahovaných hodnot příznaků od manuálně změřených hodnot a průměrná odchylka špatně extrahovaných hodnot příznaku od manuálně změřených hodnot.

Ze statistik je patrné, že pokud jsou hodnoty daného příznaku v obrázku nalezeny, jsou zpravidla úspěšně a přesně určeny. Chybu extrakce tedy častěji způsobuje nenalezení hodnoty příznaku nebo nalezení falešné hodnoty příznaku. Ve sloupci Počet snímků tabulek 5.3 a 5.4 je počet snímků, ve kterých jsou přítomny nějaké hodnoty daného příznaku, ve sloupci Správný počet je podíl snímků, kde byl nalezen správný počet hodnot, ve sloupci Falešných / Nedetekovaných je podíl nalezených falešných hodnot příznaků a podíl nedetekovaných hodnot příznaků z celkového počtu hodnot příznaků. Je vidět, že nástroj má většinou spíše tendenci detekovat falešné hodnoty příznaků než hodnoty příznaků vynechávat. Další zlepšení detekce je tedy možno směřovat směrem k snížení počtu detekovaných falešných hodnot příznaků.

• **Extrakce vertikálních souřadnic otevření a uzavření**

Extrakce vertikální souřadnice počátku a konce otevření hlasivky může selhat, pokud dochází k velmi krátkému uzavření hlasivky. Konec otevření v tomto případě není detekován a toto otevření je tak spojeno s následujícím otevřením. Místo uzavření mezi těmito dvěma otevřeními je detekován

Charakteristika	Počet	Úspěšnost v %	Průměrná chyba
Vert. souř. otevření	35 / 297	78 / 88.9 / 99.2	1.1 / 9
Vert. souř. uzavření	35 / 293	78 / 89.4 / 100	0.5 / 0
Horiz. souř. otevření	35 / 254	98 / 99.2 / 99.2	1.1 / 7.5
Horiz. souř. uzavření	35 / 254	94 / 98.8 / 98.8	1.0 / 4.3
Levé střed. vrcholy	21 / 145	84 / 86.9 / 99.2	2.2 / 19
Pravé střed. vrcholy	18 / 146	76 / 71.9 / 98.1	1.5 / 20
Horiz. souř. amplitud levé hlasivky	49 / 298	78 / 89.3 / 96.4	1 / 7.5
Horiz. souř. amplitud pravé hlasivky	47 / 288	66 / 84.7 / 94.9	1.2 / 6.2
Vertik. souř. amplitud levé hlasivky	49 / 298	68 / 89.6 / 96.7	1.5 / 11
Vert. souř. amplitud pravé hlasivky	47 / 289	64 / 82.4 / 91.9	1.6 / 11

Tabulka 5.1: Vyhodnocení úspěšnosti jednotlivé extrakce základních příznaků. Ve sloupci Počet obou tabulek je počet snímků na kterých se vyskytovala alespoň jedna hodnota daného příznaku a celkový počet hodnot tohoto příznaku na všech snímcích. Ve sloupci Úspěšnost je podíl snímků bez chybně detekované hodnoty z celkového počtu snímků, podíl všech správně detekovaných hodnot z celkového počtu hodnot a podíl všech správně detekovaných hodnot na snímcích, kde byl detekovaný správný počet hodnot daného příznaku. Ve sloupci Průměrná odchylka je průměrná odchylka správně extrahovaných hodnot příznaků od manuálně změřených hodnot a průměrná odchylka špatně extrahovaných hodnot příznaku od manuálně změřených hodnot.

Charakteristika	Počet	Úspěšnost v %	Průměrná chyba
Vert. souř. otevření	35 / 290	84 / 90 / 99.6	1.1 / 10
Vert. souř. uzavření	35 / 286	84 / 92 / 100	0.5 / 0.0
Horiz. souř. otevření	35 / 273	82 / 90.5 / 98	1.1 / 5.0
Horiz. souř. uzavření	35 / 273	76 / 90.8 / 98	1.1 / 4.0
Levé střed. vrcholy	21 / 138	86 / 87 / 98.4	2.5 / 13
Pravé střed. vrcholy	18 / 135	76 / 73.3 / 97.1	1.6 / 66.3
Horiz. souř. amplitud levé hlasivky	49 / 301	80 / 92.7 / 97.6	1.0 / 7.3
Horiz. souř. amplitud pravé hlasivky	47 / 296	72 / 87.5 / 94.9	1.2 / 5.8
Vertik. souř. amplitud levé hlasivky	49 / 301	72 / 91.7 / 96.8	1.4 / 11.3
Vert. souř. amplitud pravé hlasivky	47 / 298	60 / 82.9 / 94.6	1.5 / 9.4

Tabulka 5.2: Vyhodnocení úspěšnosti plně automatické extrakce základních příznaků. Ve sloupci Počet obou tabulek je počet snímků na kterých se vyskytovala alespoň jedna hodnota daného příznaku a celkový počet hodnot tohoto příznaku na všech snímcích. Ve sloupci Úspěšnost je podíl snímků bez chybně detekované hodnoty z celkového počtu snímků, podíl všech správně detekovaných hodnot z celkového počtu hodnot a podíl všech správně detekovaných hodnot na snímcích, kde byl detekovaný správný počet hodnot daného příznaku. Ve sloupci Průměrná odchylka je průměrná odchylka správně extrahovaných hodnot příznaků od manuálně změřených hodnot a průměrná odchylka špatně extrahovaných hodnot příznaku od manuálně změřených hodnot.

Charakteristika	Správný počet %	Falešných / Nedetekovaných %
Vert. souř. otevření	78	0.7 / 18
Vert. souř. uzavření	78	0.7 / 18
Horiz. souř. otevření	100	0 / 0
Horiz. souř. uzavření	100	0 / 0
Levé střed. vrcholy	86	2.1 / 8
Pravé střed. vrcholy	76	2.7 / 18
Horiz. souř. amplitud levé hlasivky	84	6 / 2
Horiz. souř. amplitud pravé hlasivky	82	10.4 / 2
Vertik. souř. amplitud levé hlasivky	84	6 / 2
Vert. souř. amplitud pravé hlasivky	82	10 / 2

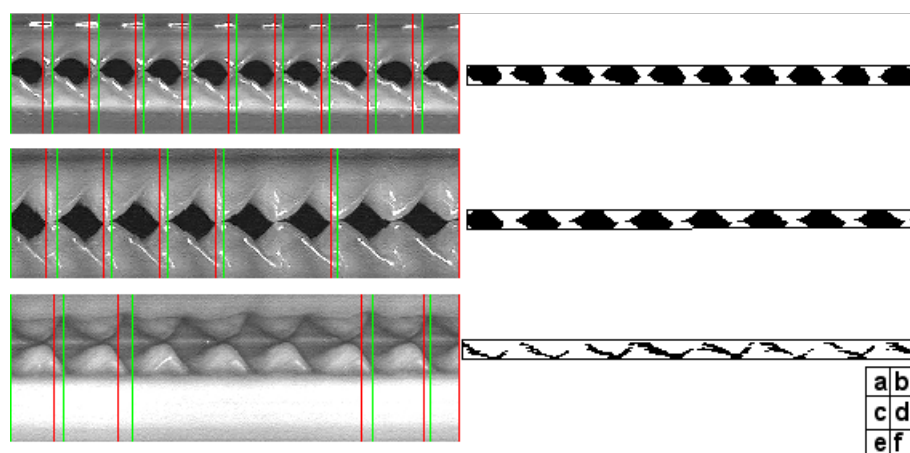
Tabulka 5.3: Další statistiky vyhodnocení jednotlivé extrakce základních příznaků. Ve sloupci Počet snímků je počet snímků, ve kterých jsou přítomny nějaké hodnoty daného příznaku, ve sloupci Správný počet je podíl snímků, kde byl nalezen správný počet hodnot, ve sloupci Falešných / Nedetekovaných je podíl nalezených falešných hodnot příznaků a podíl nedetekovaných hodnot příznaků z celkového počtu hodnot příznaků.

Charakteristika	Správný počet %	Falešných / Nedetekovaných %
Vert. souř. otevření	84	9.3 / 0.3
Vert. souř. uzavření	84	7.7 / 0.3
Horiz. souř. otevření	86	7.3 / 0.4
Horiz. souř. uzavření	86	7 / 0.4
Levé střed. vrcholy	88	6.5 / 5.1
Pravé střed. vrcholy	78	17.8 / 6.7
Horiz. souř. amplitud levé hlasivky	86	2.3 / 2.7
Horiz. souř. amplitud pravé hlasivky	82	4.1 / 3.7
Vertik. souř. amplitud levé hlasivky	86	2.3 / 3
Vert. souř. amplitud pravé hlasivky	78	4.4 / 8.1

Tabulka 5.4: Další statistiky vyhodnocení plně automatické extrakce základních příznaků. Ve sloupci Počet snímků je počet snímků, ve kterých jsou přítomny nějaké hodnoty daného příznaku, ve sloupci Správný počet je podíl snímků, kde byl nalezen správný počet hodnot, ve sloupci Falešných / Nedetekovaných je podíl nalezených falešných hodnot příznaků a podíl nedetekovaných hodnot příznaků z celkového počtu hodnot příznaků.

středový vrchol. Tato chyba nastává, protože příliš krátká otevření a uzavření jsou zvýšení robustnosti ignorována. V opačném případě by přítomnost šumu a artefaktů ve snímku mohla způsobit detekci falešných otevření a uzavření. Chybná extrakce otevření a uzavření může nastat také v případě málo kontrastního snímku a nehomogenity intenzity pixelů příslušejících do segmentu odpovídajícímu hlasivkové štěrbině. Na takovém snímku může dojít k selhání algoritmu hledání prahu, otevření hlasivky jsou špatně segmentovány. V důsledku toho pak dojde ke špatné extrakci začátků a konců otevření.

Na obrázku 5.2-a je videokymografický snímek se správně detekovanými otevřeními a uzavřeními. Na obrázku 5.2-b je segmentovaná část tohoto snímku v blízkosti jeho středu na základě které dojde k detekci začátků a konců otevření. Na obrázku 5.2-c, d je snímek s dvěma příliš krátkými uzavřeními a segmentovaná část použitá pro detekci začátků a konců otevření. Segmentace snímku proběhla úspěšně, konce otevření ale nebyly detekovány kvůli jejich malé délce. Na obrázku 5.2-e, f je znázorněno hledání začátků a konců otevření málo kontrastního snímku. Oblast snímku okolo středu byla špatně segmentována a byly proto nalezeny falešné uzavření.

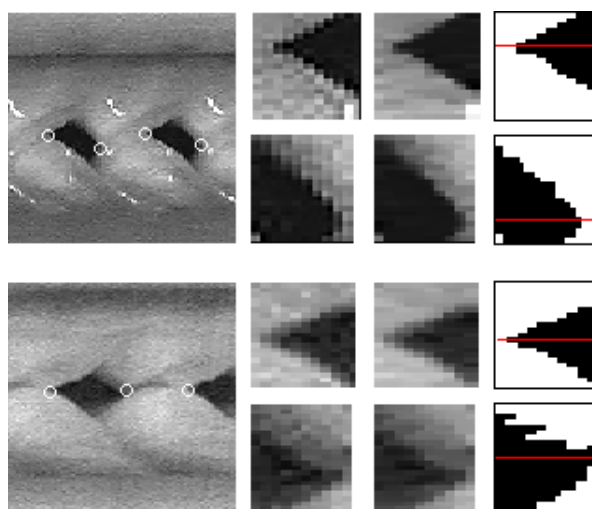


Obrázek 5.2: Extrakce vertikálních souřadnic otevření a uzavření hlasivky. Snímek se správně detekovanými začátky a konci otevření (a), segmentovaná část okolo středu tohoto snímku použitá k extrakci začátků a konců otevření, (b), snímek s chybně detekovanými uzavřeními hlasivky vlivem jejich malé délky a segmentovaná část tohoto snímku (c, d), málo kontrastní snímek s chybně detekovanými otevřeními a uzavřeními a segmentovaná část tohoto snímku (e, f).

• Extrakce horizontálních souřadnic otevření a uzavření

Chyba při extrakci horizontální souřadnice začátku a konce otevření může nastat pokud je hranice otevření rozostřená a je na ní přítomné velké množství šumu. Může pak dojít k chybné segmentaci výřezu algoritmem narůstáním oblastí.

Na obrázku 5.3 nahoře je videokymografický snímek se správně detekovanými horizontálními souřadnicemi otevření a uzavření. Vpravo jsou výřezy oblastí snímku, kde dochází k hledání zmíněných souřadnic, výřezy po filtraci šumu a segmentované oblasti začátku respektive konce otevření. Na obrázku 5.3 dole je pak snímek s rozostřenou hranicí začátku a především konce otevření a příslušné výřezy oblastí snímku. V případě začátku otevření je hranice stále dostatečně výrazná a oblast začátku otevření je segmentována správně. V případě konce otevření je rozostření větší a došlo ke špatné segmentaci oblasti. Následkem toho je horizontální souřadnice konce otevření posunutá směrem nahoru.



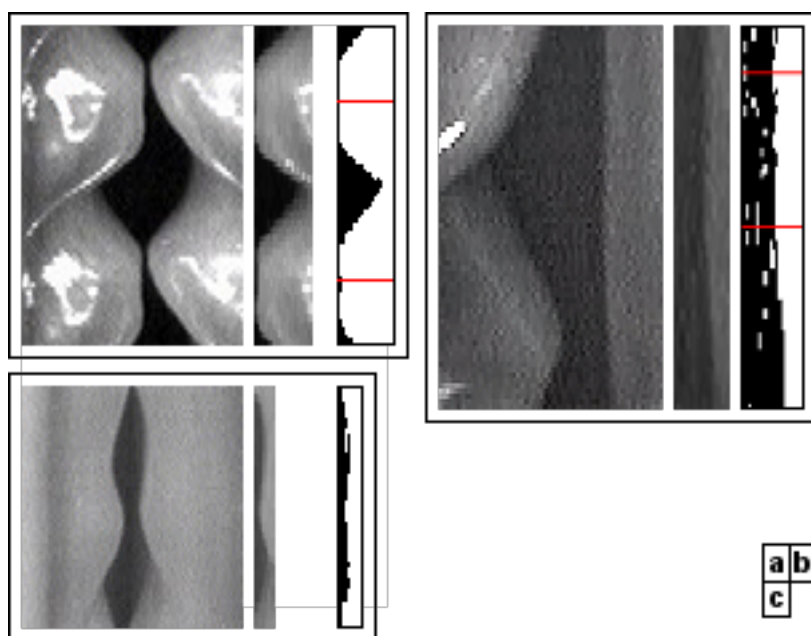
Obrázek 5.3: Extrakce horizontálních souřadnic začátku a konce otevření hlasivky. Správně detekovaný začátek i konec hlasivky, výřezy snímků, kde dochází k hledání začátku respektive konce hlasivky, tyto výřezy s potlačeným šumem a segmentovaný začátek respektive konec hlasivkové štěrbině z těchto výřezů (nahore), snímek s nesprávně detekovaným koncem otevření hlasivky a příslušné výřezy snímků (dole).

• Extrakce vertikálních souřadnic vrcholů středových vrcholů

Průměrná odchylka automaticky extrahovaných poloh vrcholů středových vrcholů od manuálně naměřených je větší než u ostatních parametrů. Je to proto, že vrcholy středových vrcholů jsou často velmi zaoblené a jejich správná poloha není někdy zcela zřejmá. Pokud snímek neobsahuje středové vrcholy, ale obsahuje větší množství šumu, může dojít k detekci falešných středových vrcholů. V případě, že snímek je zašuměný nebo je hranice mezi hlasivkovou štěrbinou a jejím okolím rozostřená a středové vrcholy jsou málo výrazné, může se naopak stát, že nejsou detekovány.

Na obrázku 5.4-a je snímek s výrazně zaoblenými vrcholy středových vrcholů levé hlasivky, odšuměný výřez snímku ve kterém se hledají středové

vrcholy a segmentovaná hlasivková štěrбина z tohoto výřezu se zvýrazněnými automaticky extrahovanými polohami středových vrcholů. Na obrázku 5.4-b je silně zašuměný snímek s detekovaným falešným středovým vrcholem a na obrázku 5.4-c je snímek s nedetekovaným málo výrazným středovým vrcholem.



Obrázek 5.4: Extrakce středových vrcholů.

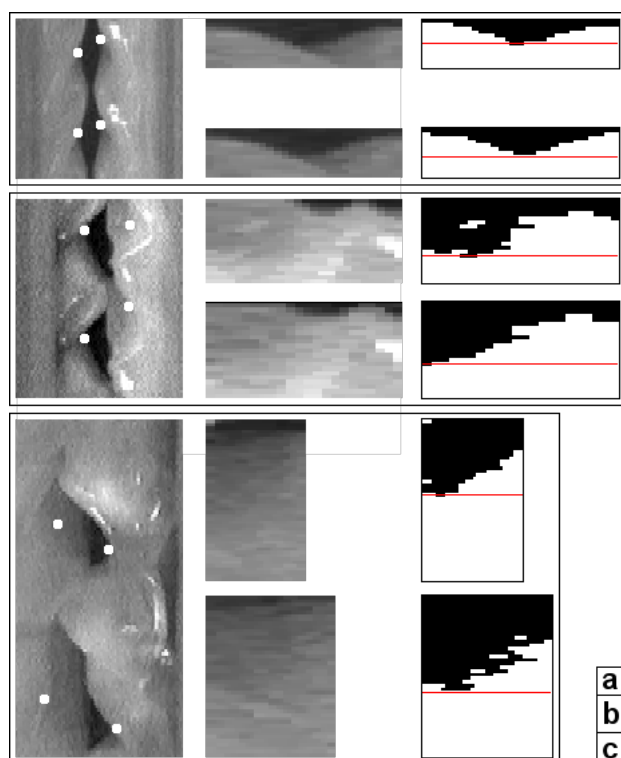
- **Extrakce horizontálních souřadnic amplitud kmitání**

Extrakce horizontálních souřadnic amplitud kmitání zpravidla selže pokud hlasivka nekmitá. Nedochází totiž k explicitní kontrole přítomnosti kmitání hlasivky a algoritmus tak vždy snaží nalézt jak amplitudy kmitů tak středové vrcholy. Tento problém by bylo možné vyřešit pomocí explicitní kontroly přítomnosti vibrací. Nabízí se kombinace kontroly periodicity signálu pomocí Fourierovy transformace, případně autokorelační funkce a kontroly přítomnosti vibrací za použití hledání otevření a uzavření hlasivky použité při hledání středu. Problém je, že vibrace hlasivek nemusí být periodické a zmíněná kontrola přítomnosti vibrací není zcela robustní - v případě hledání středu vibrací slouží pouze jako nápověda. Je tedy otázka, zda by taková kontrola nevedla k velkému počtu nedetekovaných středových vrcholů a amplitud kmitů.

Problém při segmentaci hlasivkové štěrbině a jejího okolí může nastat i pokud je amplituda kmitu příliš malá. Vlivem malého podílu pixelů hlasivkové štěrbině ve výřezu snímku, kde dochází ke hledání horizontální souřadnice amplitudy kmitání, nemusí být v histogramu patrný vrchol odpovídající této oblasti a algoritmus hledání prahu nenajde správný práh. Může se také

stát, že správně nalezený práh je na základě malého počtu pixelů tohoto výřezu menších než tento práh označen jako nesprávný. Fázi segmentace hlasivkové štěrbiny také podobně jako extrakci ostatních základních charakteristik komplikuje rozostřená hranice mezi hlasivkovou štěrbinou a jejím okolím.

Na obrázku 5.5-a je snímek se správně extrahovanými horizontálními souřadnicemi amplitud kmitání. Výřez ve kterém dochází k hledání této souřadnice není příliš zašuměný a hlasivková štěrбина je v tomto výřezu dostatečně výrazná a její segmentace tak proběhla úspěšně. Na obrázku 5.5-b je snímek s malou amplitudou kmitu pravé hlasivky. Podíl pixelů hlasivkové štěrbiny ve výřezu, kde dochází k hledání horizontální souřadnice amplitudy hlasivkové štěrbiny je příliš malý a segmentace proto nebyla úspěšná. Na obrázku 5.5-c je snímek na kterém levá hlasivka nekmitá. V příslušných výřezích téměř není hlasivková štěrбина přítomna a její segmentace tak nemůže být úspěšná.

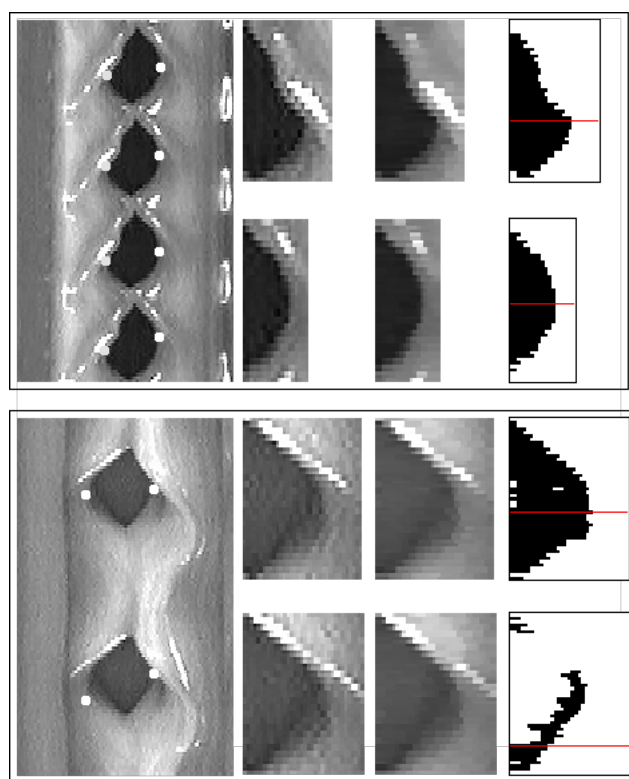


Obrázek 5.5: Extrakce horizontálních souřadnic amplitud kmitání. Snímek se správně detekovanými horizontálními souřadnicemi amplitud kmitání, výřezy snímku, kde jsou souřadnice hledány, výřezy se segmentovanou hlasivkovou štěrbinou a nalezenou horizontální souřadnicí amplitudy kmitu (a), snímek s malou amplitudou kmitů a příslušné výřezy (b), snímek na kterém levá hlasivka nekmitá a příslušné výřezy.

- **Extrakce vertikálních souřadnic amplitud kmitání**

Příčiny chybné extrakce vertikálních souřadnic amplitud jsou zpravidla stejné jako v případě extrakce horizontálních souřadnic amplitud kmitání.

Na obrázku 5.6-a je snímek se správně extrahovanými vertikálními souřadnicemi amplitud kmitání. Na obrázku 5.6-b je snímek, kde tyto souřadnice levé hlasivky nebyly extrahovány přesně. Při hledání vertikální souřadnice první amplitudy levé hlasivky došlo kvůli neostré hranici mezi hlasivkovou štěrbinou a jejím okolím k nepřesné segmentaci hlasivkové štěrbiny. Při hledání vertikální souřadnice druhé amplitudy levé hlasivky došlo k chybnému nalezení prahu vlivem malého kontrastu snímku a nehomogenity intenzity pixelů příslušející do segmentu odpovídající hlasivkové štěrbině.



Obrázek 5.6: Extrakce vertikálních souřadnic amplitud kmitání. Snímek se správně extrahovanými vertikálními souřadnicemi amplitud kmitání, výřezy snímku, kde dochází k hledání vertikální souřadnice amplitudy kmitání levé a pravé hlasivky, tento výřez po filtraci šumu a segmentovaná hlasivková štěrbina z tohoto výřezu (nahore). Málo kontrastní snímek s velkým množstvím šumu a neostrými hranicemi mezi hlasivkovou štěrbinou a jejím okolím a příslušné výřezy použité při hledání vertikálních souřadnic amplitud levé hlasivky.

Kapitola 6

Závěr a otevřené problémy

Hlavního cíle diplomové práce bylo dosaženo - byly vyvinuty algoritmy detekující základní charakteristiky patrné z videokymografických snímků. Tyto algoritmy byly popsány a otestovány na reálných datech. Flexibilní návrh celého systému a vytvořený systém pro vyhodnocení výsledků umožňuje experimentování s novými algoritmy a jejich vzájemné porovnání.

Mezi otevřené problémy patří především detekce dalších možných charakteristik kmitání. Další vylepšení mohou zahrnovat přizpůsobení nástroje k praktickému použití při vyšetření hlasivek nebo při zkoumání charakteristik jejich vibrací nebo optimalizace stávajících algoritmů. Je také možné věnovat se zvýšení přesnosti extrakce základních příznaků. Prostor ke zlepšení je zejména v zabudování explicitní kontroly přítomnosti kmitání hlasivky. Nabízí se kombinace kontroly periodicity signálu pomocí Fourierovy transformace, případně autokorelační funkce a kontroly přítomnosti vibrací za použití hledání otevření a uzavření hlasivky použité při hledání středu. Je přitom nutné vzít na zřetel to, že kmitání hlasivek nemusí být periodické.

Otevřený problém je také statistická detekce poruchy hlasivek. Ta umožňuje na základě charakteristik z videokymografického snímku statisticky určit přímo poruchu hlasivek. Pro dosažení tohoto cíle by bylo nutné podrobněji prozkoumat vibrační rysy hlasivek a shromáždit dostatečné množství statistických dat o souvislosti charakteristik vibrací hlasivek a poruch hlasivek.

Dodatek A

Obsah CD

Na přiloženém CD se nachází následující soubory:

- doc/diplomova_prace.ps ... elektronická verze diplomové práce vhodná pro tisk
- doc/diplomova_prace.pdf ... elektronická verze diplomové práce s odkazy vhodná pro prohlížení
- program/program.zip ... implementace automatické extrakce videokymografických snímků
- img/ ... sada testovacích videokymografických snímků

Dodatek B

Uživatelská příručka

B.1 Instalace

Distribuce programu obsahuje zdrojové kódy, sadu testovacích snímků a známé hodnoty příznaků těchto snímků, které byly získány manuálním měřením.

Systémové požadavky

Program pro svůj běh potřebuje prostředí Matlab [9] verze r2008a a novější. Program byl testován ve verzi tohoto prostředí r2008b.

Instalace

Pro instalaci programu je třeba zkopírovat rozbalit archiv obsahující distribuci programu do libovolného adresáře a dále zajistit oprávnění zápisu do podadresáře `processing_results/`. V případě spouštění vyhodnocení je dále třeba zajistit oprávnění zápisu do podadresáře `processing_results_evaluation/`.

B.2 Spuštění programu

Pro spuštění programu je třeba spustit prostředí Matlab, jako pracovní adresář nastavit adresář s rozbalenou distribucí programu a v konzoly programu Matlab spustit následující příkaz:

```
$ Processing[.parametr]
```

Parametr `parametr` může být následujícího tvaru:

```
runVisualisation(snimek, co_vizualizovat)
```

 pokud není argument `co_vizualizovat` zadán, spustí extrakci základních příznaků na snímku s názvem `snimek` uloženém v adresáři `input_images/`, výsledky extrakce zapíše do adresáře `processing_results/estimation/` a následně spustí jejich vizualizaci. Pokud je argument `co_vizualizovat` zadán, spustí vizualizaci daného algoritmu programu. Všechny možné hodnoty parametru `co_vizualizovat` jsou popsány v dokumentaci třídy `Processing`.

`runEstimationOfParameter(verze, priznak)` spustí extrakci příznaku daného parametrem `priznak` a výsledky extrakce spolu s konfigurací použitých algoritmů uloží do adresáře `processing_results/verze/`. Dále spustí vyhodnocení tohoto příznaku a výsledky vyhodnocení uloží do adresáře `processing_results_evaluation/verze/`. Pokud není argument `priznak` zadán, spustí extrakci všech příznaků. Všechny možné hodnoty argumentu `priznak` jsou popsány v dokumentaci třídy `Processing`.

`runEvaluation(verze)` spustí vyhodnocení dané verze extrakce příznaků uložené v adresáři `processing_results_evaluation/verze/`. Pokud není argument `verze` zadán, spustí vyhodnocení všech verzí extrakce uložených v adresáři `processing_results/`.

Spuštěním následujícího příkazu dojde k výpisu těchto parametrů do konzole spolu s jejich podrobnějším popisem:

```
$ help Processing
```

B.3 Vizualizace příznaků ve snímku

Po zadání příkazu `idProcessing.runVisualisation(snimek)` se nejprve spustí extrakce všech základních příznaků ze snímku, extrahované hodnoty a konfigurace algoritmů extrakce se zapíšou do souborů v adresáři `processing_results/estimation/`, určí se hodnoty odvozených příznaků a příznaky se zobrazí. Dojde k otevření tří oken. V prvním okně je zobrazen celý snímek hlasivky s vyznačenými polohami středu snímku, výřezu okolo středu snímku, kde dochází k hledání dalších základních příznaků, vertikálních a horizontálních souřadnic začátků a konců otevření a amplitud kmitů. Amplitudy kmitů se hledají pouze v rámci míst amplitud, které jsou zcela kompletní, tedy nesousedí se začátkem nebo koncem snímku. Dále je vypsán počet extrahovaných otevření a uzavření hlasivky, délky otevření a uzavření, poměr mezi průměrnou délkou otevření a uzavření, horizontální posunutí os mezi konci otevření a následujícími začátky otevření a průměrná hodnota tohoto posunutí.

Ve druhém a třetím okně jsou zobrazeny části videokymografického snímku odpovídající kmitání levé respektive pravé hlasivky. Na nich jsou vyznačeny horizontální souřadnice amplitud kmitů a vertikální souřadnice středových vrcholů. Je vypsán počet extrahovaných amplitud kmitání, počet extrahovaných středových vrcholů, dále poměry mezi délkou otevírání a zavírání hlasivky a průměrná hodnota těchto poměrů. Jsou také vypsány hodnoty amplitud hlasivky a průměrná hodnota amplitudy, délky period kmitání hlasivky a průměrná hodnota periody.

Příklad vizualizace příznaků ve snímku je na obrázku B.1.

Typ příznaků	Jméno souboru
Vertikální souřadnice začátků otevření	<code>openings_starts_vertical_coordinates</code>
Vertikální souřadnice konců otevření	<code>openings_ends_vertical_coordinates</code>
Horizontální souřadnice začátků otevření	<code>openings_starts_horizontal_coordinates</code>
Horizontální souřadnice konců otevření	<code>openings_ends_horizontal_coordinates</code>
Vertikální souřadnice středových vrcholů levé hlasivky	<code>medial_peaks_left</code>
Vertikální souřadnice středových vrcholů pravé hlasivky	<code>medial_peaks_right</code>
Horizontální souřadnice amplitud levé hlasivky	<code>amplitudes_values_left</code>
Horizontální souřadnice amplitud pravé hlasivky	<code>amplitudes_values_right</code>
Vertikální souřadnice amplitud levé hlasivky	<code>amplitudes_locations_left</code>
Vertikální souřadnice amplitud pravé hlasivky	<code>amplitudes_locations_right</code>

Tabulka B.1: Názvy souborů obsahující výchozí a známé hodnoty příznaků daného typu.

B.4 Zadání výchozích a známých hodnot základních příznaků

Pokud jsou zadány výchozí hodnoty základních příznaků, nebude program dané příznaky automaticky extrahovat, ale použije tyto zadané výchozí hodnoty. Známé hodnoty základních příznaků slouží k vyhodnocení automatické extrakce základních příznaků. Výchozí hodnoty základních příznaků jsou uloženy v adresáři `default_parameter_values`, známé hodnoty v adresáři `parameter_values`. Oba tyto adresáře mohou obsahovat textové soubory s hodnotami výchozích respektive základních příznaků. Seznam textových souborů, které se v těchto adresářích mohou nacházet, je v tabulce B.1.

Pro zadání výchozí respektive známé hodnoty daného základního příznaku je třeba otevřít příslušný soubor, případně ho vytvořit, pokud neexistuje. Do tohoto souboru je dále nutné přidat řádek následujícího formátu:

`jméno obrázku; hodnoty příznaku`

Parametr `hodnoty příznaku` může být číslo, seznam čísel oddělený mezerou nebo hodnota `none`. Tato hodnota znamená absenci příznaku ve snímku. Pokud je parametr `hodnoty příznaku` prázdný, příslušná řádka se ignoruje. V souboru je možné používat řádkové komentáře. Řádka začínající znakem `#` je při zpracování souboru ignorována.

Jméno souboru / adresáře	Popis
<code>badly_estimated/</code>	Obsahuje snímky se špatně určenou hodnotou daného základního příznaku.
<code>not_evaluated/</code>	Obsahuje snímky, jejichž hodnoty daného základního příznaku byly extrahovány, ale nebyly zadány známé hodnoty a automaticky extrahované hodnoty tedy nemohly být vyhodnoceny.
<code>well_estimated</code>	Obsahuje snímky se správně určenou hodnotou základního příznaku.
<code>jmeno_příznaku_badly_estimated</code>	Obsahuje seznam obrázků se špatně určenou hodnotou daného základního příznaku.
<code>jmeno_příznaku_not_evaluated</code>	Obsahuje seznam snímků, jejichž hodnoty daného základního příznaku byly extrahovány, ale nebyly zadány známé hodnoty a automaticky extrahované hodnoty tedy nemohly být vyhodnoceny.
<code>jmeno_příznaku_differences</code>	Pro každý obrázek a každou automaticky extrahovanou hodnotu příznaku obsahuje její odchylku od známé hodnoty příznaku.
<code>jmeno_příznaku_distances</code>	Pro každý obrázek obsahuje nejvyšší odchylku automaticky extrahované hodnoty od známé hodnoty příznaku.
<code>jmeno_příznaku_overall_statistics</code>	Obsahuje statistiky vyhodnocení automatické extrakce daného příznaku.

Tabulka B.2: Obsah adresáře obsahující vyhodnocení daného příznaku.

B.5 Vyhodnocení výsledků

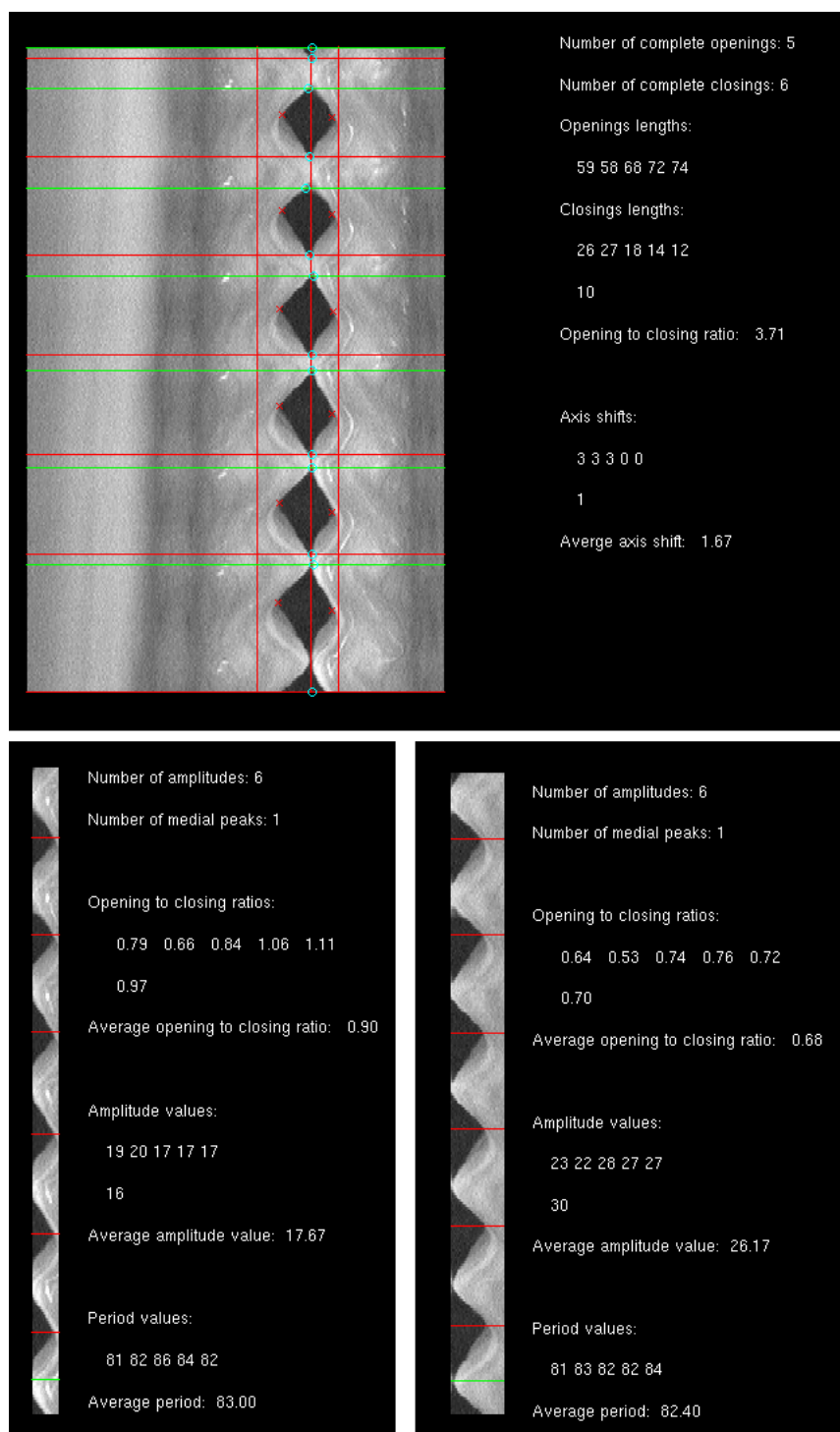
Příkaz `runEvaluation(verze)` spustí vyhodnocení výsledků automatické extrakce základních příznaků specifikované verze a do adresáře `processing_results_evaluation/verze/` vyhodnocení zapíše. V tomto adresáři se pro každý vyhodnocený základní příznak nachází adresář s obsahem popsaným v tabulce B.2.

Odchylka automaticky extrahované hodnoty od známé hodnoty může nabývat několika speciálních hodnot. Hodnoty příznaku, které byly nalezeny ve snímku, kde se daný příznak nevyskytuje jsou označeny hodnotou 1001, naopak pokud se příznak ve snímku vyskytuje a žádná hodnota nebyla detekována, jsou známé hodnoty označeny hodnotou 1002. Hodnotou 1004 jsou označeny hodnoty, které byly detekovány navíc ve snímku, kde se daný příznak vyskytuje a hodnotou 1003 hodnoty, které nebyly detekovány ve snímku, kde byla detekována přítomnost daného příznaku. Dále hodnota 0.01 nejvyšší odchylky automaticky extrahované hodnoty od známé hodnoty příznaku pro daný obrázek značí, že se ve snímku daný příznak nevyskytuje a nebyl detekován.

Příznak ve snímku je považován za špatně extrahovaný pokud je nejvyšší odchylka automaticky extrahované hodnoty od známé hodnoty tohoto příznaku větší než daná tolerance. Toleranci lze pro každý příznak nastavit editací příslušných souborů adresáři `parameter_values/additional/`. Seznam statistik v souboru `jméno_příznaku_overall_statistics` je uveden v tabulce B.3.

Název statistiky	Popis
Number of evaluated images	Počet vyhodnocovaných snímků.
Number of badly estimated images	Počet snímků na kterých byla některá hodnota daného příznaku špatně určena.
Number of well estimated images	Počet snímků na kterých byly správně určeny všechny hodnoty daného příznaku.
Portion of well estimated images	Podíl snímků na kterých byly správně určeny všechny hodnoty daného příznaku z celkového počtu snímků.
Number of images where some parameter value is present	Počet snímků na kterých je daný příznak přítomný.
Portion of images where it was detected correct number of values	Podíl snímků, kde byl detekován správný počet hodnot příznaku.
Portion of images with less values detected	Podíl snímků, kde bylo detekováno méně hodnot.
Portion of images with more values detected	Podíl snímků, kde bylo detekováno více hodnot.
Detected absence of parameter value in images where no parameter value was present	Podíl snímků ve kterých byla detekována absence příznaku ze všech snímků daný příznak neobsahující.
Detected presence of parameter value in images where some parameter value was present	Podíl snímků, kde byla detekována přítomnost příznaku ve všech snímcích hodnotu příznaku obsahujících.
Number of parameter values	Celkový počet hodnot příznaku ve všech obrázcích.
Number of well estimated values	Počet správně určených hodnot příznaku.
Number of badly estimated values	Počet špatně určených hodnot příznaku.
Portion of well estimated values	Podíl dobře určených hodnot z celkového počtu hodnot příznaku.
Portion of well estimated values from images where correct number of parameters was found	Podíl dobře extrahovaných hodnot příznaku v obrázcích, kde byl nalezen správný počet hodnot tohoto příznaku.
Portion of false values	Podíl falešných hodnot.
Portion of not detected values	Podíl nedetekovaných hodnot.
Average error on well estimated values with parameter value present	Průměrná chyba u správně extrahovaných hodnot příznaků.
Average error on badly estimated values	Průměrná chyba u špatně extrahovaných hodnot příznaků.

Tabulka B.3: Seznam statistik v souboru jméno_příznaku_overall_statistics.



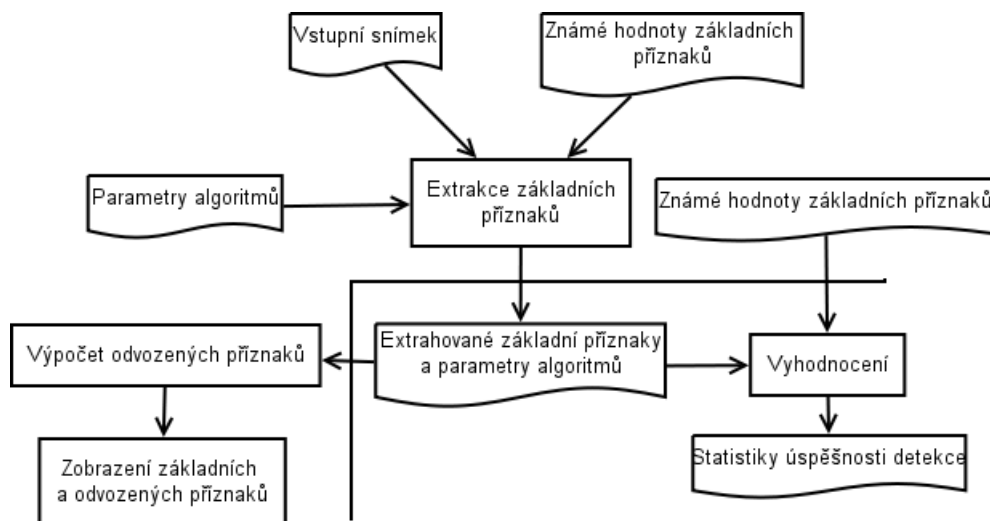
Obrázek B.1: Vizualizace příznaků ve snímku.

Dodatek C

Architektura programu

Návrh programu vychází z principů objektově orientovaného programování. Implementace byla vytvořena v prostředí Matlab [9] za použití objektově orientované syntaxe představené ve verzi r2008a. Architektura programu je popsána pomocí následujících diagramů.

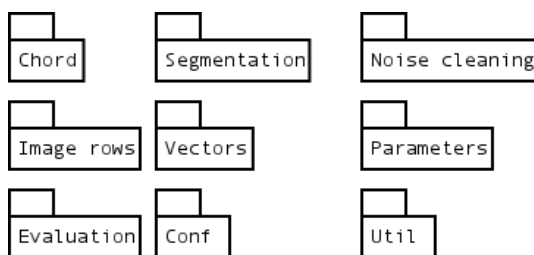
Na obrázku C.1 je tok dat v programu. Ze vstupního snímku se extrahují základní příznaky. Při extrakci se použijí parametry algoritmů zadané v konfiguraci programu a případně také zadané také známé hodnoty některých základních příznaků. Tyto základní příznaky se spolu s použitými parametry algoritmů extrakce zapíší do souboru. Na základě základních příznaků je možné spočítat další, odvozené příznaky a spolu se základními příznaky je zobrazit. Alternativně lze hodnoty extrahovaných příznaků vyhodnotit oproti známým hodnotám těchto příznaků a zapsat vyhodnocení do souboru.



Obrázek C.1: Tok dat v programu.

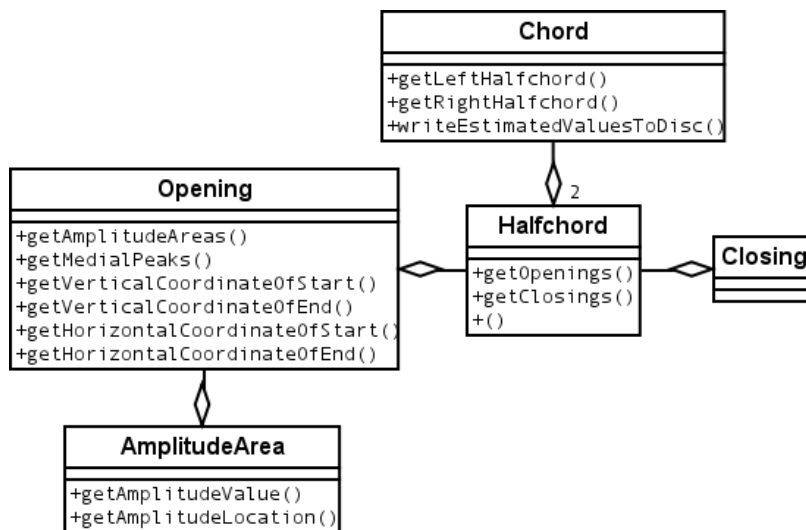
Na obrázku C.2 jsou moduly programu. Modul **Chord** obsahuje třídy reprezentující části videokymografického snímku. Modul **Segmentation** poskytuje algoritmy pro segmentaci obrazu. Modul **Noise cleaning** poskytuje algoritmy pro filtraci šumu v obrazu. Modul **Image rows** poskytuje prostředky pro výběr důleži-

tých řádků ze snímku. Tento modul obsahuje algoritmus popsáný v části 4.2. Modul **Vectors** obsahuje prostředky pro práci s vektory a pro výběr hodnot z vektoru. Modul **Parameters** umožňuje pracovat se známými a automaticky určenými hodnotami příznaků. Modul **Evaluation** slouží k vyhodnocování úspěšnosti extrakce základních příznaků. Modul **Conf** obsahuje prostředky pro konfiguraci použitých algoritmů a zápis této konfigurace do souboru. Modul **Util** obsahuje pomocné funkce a datové struktury.



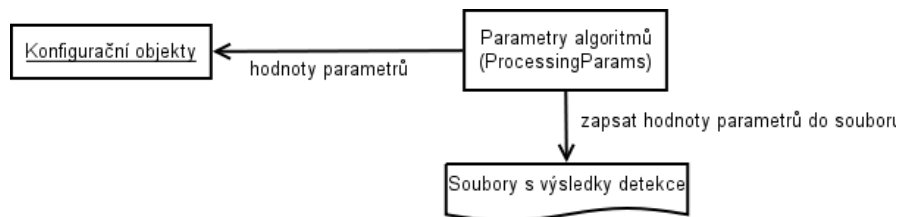
Obrázek C.2: Moduly programu.

Na obrázku C.3 jsou třídy reprezentující části videokymografického snímku. Třída **Chord** reprezentuje oblast kmitů obou hlasivek ve snímku. Třída **Halfchord** reprezentuje polovinu této oblasti obsahující kmitu jedné z hlasivek. Třída **Opening** reprezentuje oblast otevření hlasivky, třída **Closing** oblast jednoho uzavření hlasivky. Třída **AmplitudeArea** reprezentuje oblast obsahující jednu amplitudu kmitu.

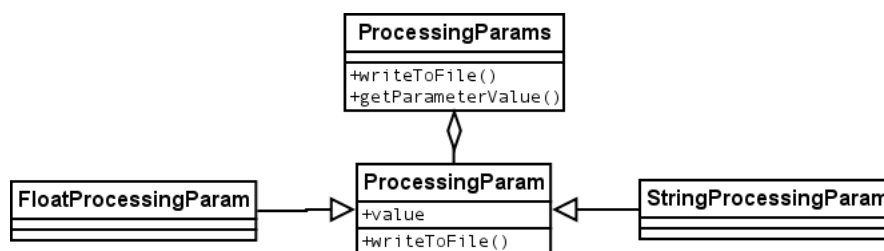


Obrázek C.3: Třídy reprezentující části videokymografického snímku.

Na obrázku C.4 je zobrazen tok dat při práci s parametry algoritmů. Statická třída **ProcessingParams** slouží jako kontejner pro hodnoty parametrů algoritmů. Umožňuje zapsat tyto parametry do souboru a vrátit hodnotu daného parametru. Konfigurační objektu tuto statickou třídu používají pro získání hodnot parametrů algoritmů. Vlastní hodnoty parametrů algoritmů obsahují instance třídy **ProcessingParam** znázorněné na obrázku C.5.

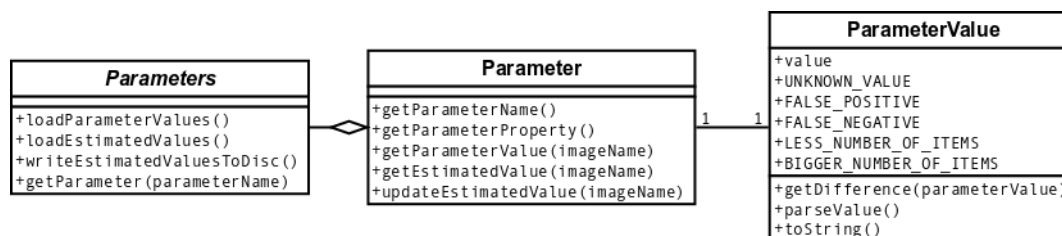


Obrázek C.4: Tok dat při práci s parametry algoritmů.



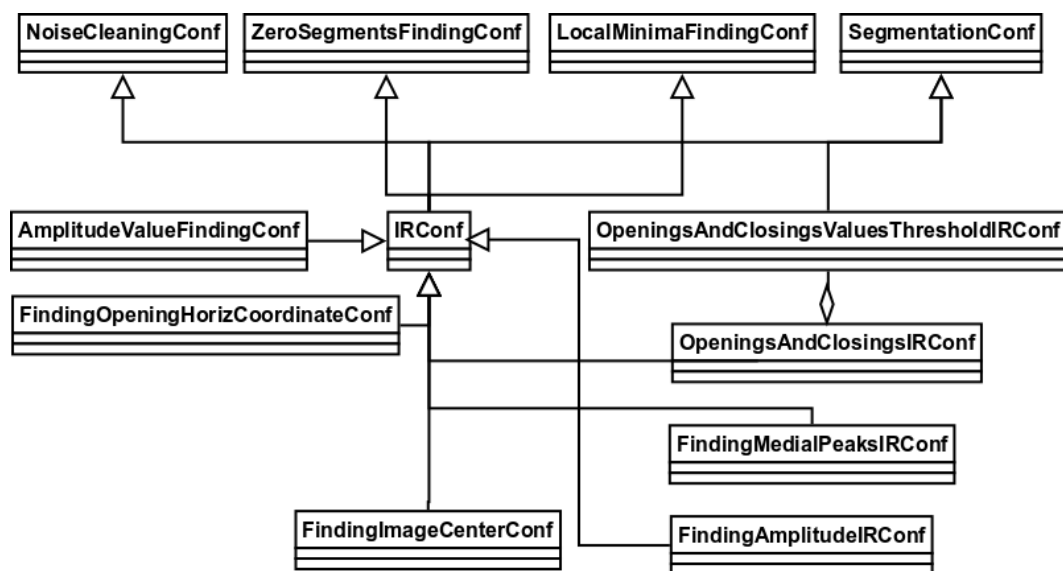
Obrázek C.5: Třídy reprezentující parametry algoritmů.

Na obrázku C.6 jsou znázorněny třídy modulu `Parameters`. Instance třídy `Parameter` uchovávají známé a automaticky extrahované hodnoty daného příznaku. Singleton třídy `Parameters` v sobě obsahuje instance třídy `Parameter` obsahující hodnoty všech základních příznaků. Třída `ParameterValue` reprezentuje hodnoty daného příznaku v jednom videokymografickém snímku.



Obrázek C.6: Třídy umožňující práci s hodnotami základních příznaků.

Na obrázku C.7 je zachycená hierarchie konfiguračních tříd. Instance konfigurační třídy `NoiseCleaningConf` slouží pro konfiguraci filtrace šumu, instance třídy `SegmentationConf` pro konfiguraci segmentace, instance třídy `LocalMinimaFindingConf` pro konfiguraci hledání lokálních minim ve vektoru hodnot a instance třídy `ZeroSegmentsConf` slouží pro konfiguraci hledání nulových úseků v daném vektoru. Instance třídy `IRConf` slouží pro konfiguraci algoritmu pro výběr řádků ze snímku popsany v části 4.2. Instance dalších uvedených konfiguračních tříd slouží pro konfiguraci algoritmů hledání konkrétních základních příznaků ve snímku.



Obrázek C.7: Hierarchie konfiguračních tříd.

Literatura

- [1] prof. PhDr. Marie Krčmová, CSc.: Fonetika a fonologie,
<http://is.muni.cz/do/1499/el/estud/ff/js08/fonetika/ucebnice/index.html>
- [2] Mihael Kass, Andrew Witkin, Demetri Terzopoulos (1987): *Snakes: Active Contour Models. International Journal of Computer Vision*, 321-331 (1988).
- [3] Švec JG, Schutte HK (1996): *Videokymography: high-speed line scanning of vocal fold vibration. J Voice*. 10, 201-205.
- [4] Švec JG, (2007): *Video Kymography Evaluation Sheet*.
- [5] Dunker E, Schlosshauer B.: *Irregularities of the laryngeal vibratory pattern in healthy and hoarse persons. In: Brewer DW, ed. Research potentials in voice physiology. Syracuse, NY: State University of New York, 1964: 151-84.*
- [6] Švec JG, Šram F, Schutte HK (2007): *Videokymography in voice disorders: what to look for? Ann Otol Rhinol Laryngol*. 116, 172-180.
- [7] Rafael C. Gonzalez, Richard E. Woods: *Digital Image Processing*, PEARSON Prentice Hall, 2008
- [8] Vetešník A (2004): *Digital Image Processing of Videokymography Recordings. Project progress report*.
- [9] Matlab: high-level language and interactive environment,
<http://www.matlab.com/>